NMR Dynamics Studies of Rho130-Polymer Conjugates and Human Thymidylate Kinase

Ian James Fucci

Submitted in partial fulfillment of the requirements for the degree
of
Doctor of Philosophy
at
Carnegie Mellon University
Department of Biological Sciences
Pittsburgh, Pennsylvania

Advised by Dr. Gordon S. Rule

December 16, 2019

To my loving wife and best friend, Erika...

Acknowledgements

I would like to acknowledge the many people who made this work possible. First I would like to thank my advisor Dr. Gordon Rule, who from half a world away managed to provide continuous guidance and support. His patience and kindness have been greatly appreciated throughout the years. I would also like to thank the members of my research committee: Dr. David Hackney, Dr. Jonathan Minden and Dr. Rieko Ishima for their helpful comments and advice.

I would like to thank the many people across different universities who allowed us to use their instruments. At CMU, Virgil Simplaceanu and Roberto Gil were incredibly helpful when setting up experiments. I would also like to thank Mike Delk from Pitt and Tapas Mal from Penn State.

I am truly grateful to have made so many friends along my journey. I would like to thank members of the Rule Lab past and present: Dr. Andy Kehr, Dr. John Pettersson, Dr. Sahil Sangani, Paul Barton, Dr. David Chen and Dr. Kaustubh Sinha. It has truly been a pleasure to work with you all. I had the privilege to work on a number of projects with Kaustubh and have learned so much along the way. David and I joined the lab at the same time, and he has become a great friend over the years. I would also like to thank Dr. Amber Lucas, Dr. Stephanie Biedka, Andrew Wolff and Dan Wilson who have become some of my closest friends. Game night was always a blast.

I would like to thank my family for their unwavering support for my decision

to pursue a Ph.D. To my mother Elaine, father Paul and two brothers Ryan and Evan: I could not have gotten this far without you. Finally, to my wife Erika: we have grown so much in these past fourteen years together. You have been with me for more than half of my life and I wouldn't want to be with anyone else. I am eternally grateful for your love and support throughout our time in Pittsburgh, and I look forward to new adventures together.

Abstract

Nuclear magnetic resonance (NMR) studies on two systems: the RNA binding domain of the *E. coli* transcriptional termination factor Rho (Rho130), and human thymidylate kinase (hTMPK) will be discussed. Protein-polymer conjugates produced by atom transfer radical polymerization (ATRP) are an emergent tool in biomedical sciences with uses ranging from pharmaceuticals to biomaterials. Rho130 was used as a model system to explore the structural and dynamical implications of protein modification by polymers grown from the protein surface. Formation of protein-polymer conjugates is a two-step process consisting of: (1) the attachment of an ATRP initiator to the protein and (2) growth of polymers from the protein-attached initiator. By tracking chemical shift perturbations (CSPs) of ^{15}N- and $^{13}CH_3$ ILV methyl-labeled Rho130 it was found that the covalent attachment of initiators, and not subsequent polymerization, is more perturbing to protein structure. Initiator attachment results in additional resonance lines in the ^{15}N HSQC corresponding to the amide bonds between the initiator and accessible primary amines on the protein that were used to map degree of modification, which was confirmed by mass spectrometry. Methyl ^{1}H relaxation dispersion experiments revealed essentially no changes to the μs-ms dynamics of Rho130-conjugates. These findings have implications for the quality control of clinically relevant protein-polymer conjugates.

TMPKs catalyze the phosphorylation of TMP using ATP and Mg^{2+} to form TDP which is subsequently phosphorylated to form the DNA building block TTP. The activation of 3'-azido-3'-deoxythymidine (AZT), an anti-retroviral prodrug, is rate-limited by the conversion of AZT monophosphate (AZTMP) to the diphosphate by hTMPK; a reaction that occurs at a rate 70-fold slower than TMP conversion. A previously reported point mutant of hTMPK (F105Y) confers a 20-fold increase in activity toward AZTMP while reducing activity with the natural substrate (TMP) by 4-fold. The interaction of wild-type (WT) and F105Y hTMPK with both TMP and AZTMP was investigated using solution state NMR experiments, which showed that all complexes adopt the same average structure in solution and have few differences in dynamics on the ps-ns time scale. Methyl ^{1}H relaxation dispersion experiments indicate that notable changes to the μs-ms time scale may lead to differences in catalytic rate. It was found that AZTMP causes significant perturbations to WT hTMPK dynamics in both the ATP- and TMP-binding sites. This active site instability is rescued by the F105Y mutant, which reduces conformational flexibility in these regions leading to the decrease in activity with TMP. Interestingly, the frequency of the motions in the active site is most similar between F105Y hTMPK bound to AZTMP and WT bound to TMP, suggesting the increased AZTMP activity of the F105Y mutant may stem from similar active site tuning.

Contents

List of Figures

List of Tables

Chapter 1

Rho130-Polymer Introduction

Polymer-based protein engineering (PBPE) provides an orthogonal approach to the myriad ways of tuning enzyme function through traditional genetic approaches. In its simplest form, PBPE involves the attachment of a pre-synthesized polymers via reactive groups found on the protein; typically the thiol groups of cysteine or primary amines of lysine side chains and the N-terminus, though other reactive side chains have been exploited. The first such protein-polymer conjugate reported by Abuchowski et al. (1977) was bovine serum albumin (BSA) with a polyethylene glycol (PEG) moiety attached to the primary amines of via a cyanuric chloride reagent. This watershed moment in protein-based biological therapeutics was realized upon injection into rabbits which did not produce an immune response for the PEGylated BSA, as it did for the unmodified protein. Over the following four decades a total of 17 PEGylated proteins have been approved by the FDA for treatments of diseases as diverse as cancer, hepatitis and rheumatoid arthritis (FDA.gov). This multibillion dollar industry both requires alternatives to PEG for the purposes of delivery and compatibility, and a proper understanding of the implications of chemical modification of proteins and their complex interactions with the human body.

A number of approaches for conjugation of polymers onto proteins have been developed in recent years in an attempt to enter this lucrative market and provide enhancements to the unfunctionalized PEG groups. One such method of PBPE is described in Section 1.1. Briefly in this "grafting-from" technique, initiator groups are attached to the reactive primary amines of the protein, followed by *in situ* polymerization from these sites (Averick et al., 2012). This results in highly modified proteins with multiple polymers attached in differing quantities at each site. Naturally, it is likely that these changes to the molecule result in changes to structure and dynamics of the protein, especially in regions proximal to the polymer attachment points. Moreover, the presence of positional isomers and different polymer lengths at each site contribute to the complexity of atomic resolution studies of molecular structure and dynamics. Very few experimental studies have attempted to answer whether the underlying protein in these conjugates has significant perturbations to its natural behavior. In Section 1.2 a case is made for using nuclear magnetic resonance (NMR) spectroscopy as a powerful tool to investigate such heterogeneous mixtures of protein-polymer conjugates. Many studies report significant losses to enzyme activity (Gauthier and Klok, 2010), though lack much explanation other than insisting the polymers act as a steric barrier to substrates. Here, it is described which techniques can be used to determine the possible dynamical perturbations which occur after initiator attachment and subsequent polymerization using the model protein Rho130 (described in Section 1.3). The polymers used for this model conjugate are negatively charged carboxylate-functionalized methacryalates.

1.1 Polymer Based Protein Engineering

In recent years an explosion of new chemistry for PBPE has been developed.(Grover and Maynard, 2010; Qi and Chilkoti, 2015) Growing concerns

about the immunogenicity of PEG have been hotly debated in the field (Schellekens et al., 2013), fueling innovation of new PEG alternatives. Rather than focus on the many different methodologies, a discussion of the major techniques for conjugation and polymer synthesis employed in academic laboratories follows. Conjugation chemistry largely falls into two categories: "grafting-to" and "grafting-from" describing whether the polymers are synthesized prior to reaction with the protein, or if the polymers are grown from the surface of protein via initiators that have been attached prior to polymerization (Carmali et al., 2017). Production of the desired polymer has no restrictions in the grafting-to scheme, while in the grafting-from scheme only those polymerization techniques that are compatible with aqueous media and are not cross-reactive with the protein can be chosen. Grafting-to suffers greatly from the problem of positional isomers, or heterogeneous attachment of the polymer to many sites on the protein. Because of the large size of the reactant, steric effects play a role in preventing modification of some sites after a nearby one has already been modified. This can be prevented by selective attachment to the N- or C-termini or by recombinant expression of proteins with a single reactive center. Grafting-from partially solves this problem provided the protein-reactive initiator is much smaller than macromolecular polymers. However, *in situ* polymerization techniques used in the grafting-from scheme result in a distribution of polymer sizes attached to the protein, unlike the grafting-to method where polymers can be purified prior to attachment. This problem is largely mitigated in grafting-from methods by choosing polymerization chemistry with low dispersity, or distribution of products (Carmali et al., 2017). Combined methodologies have also been developed, where a short polymer (usually 5-10 monomers in length) containing a reactive center is first attached (Lucius et al., 2016). Subsequently, chain elongation is performed from this reactive group using a polymerization method of choice. This combined grafting-to

and grafting-from method is supposed to mitigate the risks of decreased water solubility of protein-initiator conjugates. Each method has its drawbacks, but all are active areas of research for different applications of protein-polymer conjugates. Protein stability is the major problem plaguing these conjugation techniques. Loss of charge at the protein surface often leads to deleterious effects in protein structure and function, though conjugation techniques to reintroduce the lost charge have been developed (Baker et al., 2019b). Once a suitable conjugation technique is found for a system though, the realm of possibilities for different polymer conjugates is seemingly endless. A subset of these possibilities is discussed in the following section.

Applications of Protein-Polymer Conjugates

The original protein-polymer conjugates were developed for the purpose of increased *in vivo* lifetime of therapeutic proteins. And while this is a major area of research in the field, it just barely scratches the surface of the potential applications of these biomaterials. The potential applications of protein-polymer conjugates are diverse; ranging from altering substrate specificity based on charge in chymotrypsin (Murata et al., 2013; Cummings et al., 2014) and size in avidin (Lucas et al., 2019; Kaupbayeva et al., 2019) to stabilization of lysozyme when lyophilized (Mancini et al., 2012), to self-assembly of a three enzyme nanoreactor (van Dongen et al., 2009). Using the combined properties of self-assembling viral proteins, and so-called "smart" polymers, it will be interesting to see what the future holds (Duncan and Vicent, 2013; Moatsou et al., 2015). A considerable number of studies have focused on stabilization of proteins to organic solvents (Cummings et al., 2016), pH, and temperature (Murata et al., 2013) and increasing protein solubility (Baker et al., 2019a). Altering the physical properties of proteins to be more amenable to industrial processes is a major driving force in this space. The efficiency and specificity

of enzymes catalysts are of great interest, and the ability to reduce aggregation and increase stability through PBPE should offer a whole new class of industrially relevant biocatalysts.

Grafting-from using Atom Transfer Radical Polymerization

In the grafting-from scheme used in this work, a two-step process results in highly functionalized protein with the polymer of choice (Figure 1.1). The first step involves attachment of a bromine-containing moiety to the primary amines of the protein via NHS-ester chemistry. As discussed in the previous section, the relatively small size of this molecule provides proteins modified at nearly all accessible primary amines, granted the NHS reagent is used in significant excess. Sub-stoichiometric modification of proteins is achievable, but will likely result in multiple positional isomers, as the chemistry is fairly non-selective. These protein-initiator conjugates are then subjected to atom transfer radical polymerization (ATRP) to further modify the surface of the protein. ATRP proceeds via a copper bromide catalyzed reaction, resulting in sequential addition of individual monomers. The copper catalyst begins in the oxidized form, Cu(II)Br_2, complexed to the ligand and is reduced by addition of sodium ascorbate in a process known as activators are generated by electron transfer (AGET). Reaction progress is split into three phases known as: initiation, propagation and termination. In the initiation step the copper bromide catalyst reacts with the initiator (R-Br in Figure 1.2) to transfer the halogen atom from the initiator (Br) oxidizing the copper catalyst (Cu(II)Br_2) and forming a radical ($R^{\cdot}$) (Jakubowski and Matyjaszewski, 2005). Activation of the copper catalyst is relatively slow compared to the rated of deactivation, resulting in a small number of radicals in solution at any given time. The reaction proceeds into the propagation phase by addition of a monomer, then is deactivated by the oxidized Cu(II)Br_2 to reform the reactants (Wang and

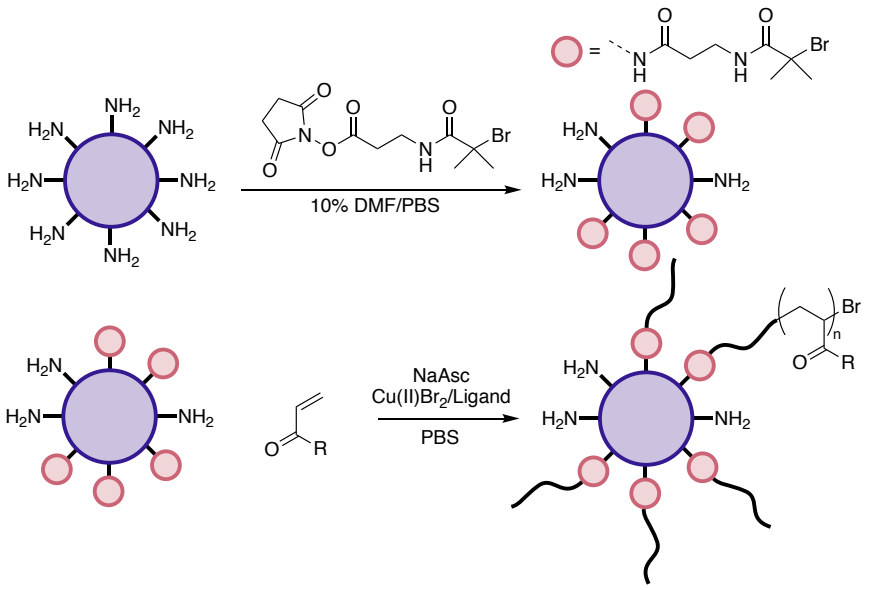


Figure 1.1: Grafting-from ATRP-based PBPE scheme. ATRP-based PBPE is split into two steps. The first attaches the ATRP initiator (represented as red circles) to the protein (violet circles). The second step is growth of polymers from the initiators (black curves) using ATRP. R represents a polymer side chain.

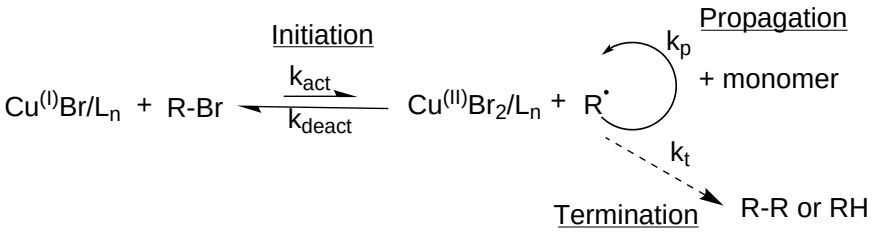


Figure 1.2: ATRP reaction mechanism. The reaction proceeds via a radical-catalyzed mechanism consisting of 3 phases: initiation, propagation and termination. L_n is the ligand, R-X is the protein-initiator complex or growing polymer chain, k_{act}, k_{deact}, k_p, and k_t are the rate constants for activation, deactivation, propagation and termination, respectively. R-R represents two active chains coming together and RH the abstraction of a proton and loss of the Br functionality on the polymer chain.

Matyjaszewski, 1995). These two phases can repeat until the monomer is consumed, providing control over the average length of the polymers by adjusting the concentration of the monomer species. Termination is a rare event due to the small concentration of radicals in solution, thus by opening the reaction to air the copper catalyst is oxidized and the reaction is stopped. This highlights a major disadvantage to ATRP for PBPE; that it must be carried out in inert conditions, requiring dissolved oxygen to be removed from protein solutions by bubbling inert gas through them. AGET ATRP is amenable to small amounts of dissolved oxygen, and techniques have been developed using the glucose oxidase/catalase system to remove excess oxygen without the need for possibly destructive bubbling (Enciso et al., 2018).

Detailed characterization of the products of the ATRP reaction has proven difficult. Typically, conjugates are subjected to analyses such as dynamic light scattering, gel permeation/size exclusion chromatography and/or mass spectrometry (Carmali et al., 2017). These techniques, while useful, provide information on the average size of the particles and approximate number of modifications, but not the interactions between the protein and polymer nor the impact of the attached polymer on protein structure and function. If the target protein is an enzyme simple binding and kinetics assays have been used to determine the impact of polymers on binding and catalysis. Studies of polymer length or protein content are performed after cleavage of the polymers from the protein in high concentrations of base (>5 M). This removes each of the components from their context in the conjugate and effectively ignores the possibility of interactions between them. In order to collect high resolution structural information on protein-polymer conjugates, a technique with high sensitivity to changes in local environment and tolerance to mixtures of analytes must be chosen. There have been limited situations where X-ray crystallography and cryoEM have provided valuable insight into structural changes to the protein

brought about by either initiator groups or attached polymers (Lucius et al., 2016). However, these two techniques do not perform well in the case of heterogeneous mixtures of protein-polymer conjugates. In the following section it will be shown that in contrast to these other techniques, NMR is an appropriate and powerful tool for studying the behavior of protein-polymer conjugates in solution.

1.2 NMR Methods for Characterizing Protein-Polymer Conjugates

Any NMR techniques that have been developed for proteins could, in principle, be used to study protein-polymer conjugates. The significant increase in size after the addition of polymers to a protein (using either grafting-from or grafting-to) is limiting because as the size of the molecule increases so do NMR line widths. At a certain point (>50 kDa) the lines in the amide ^{15}N heteronuclear single quantum coherence (HSQC) spectrum are so broad that they are indistinguishable from noise. This experiment is widely used in protein NMR spectroscopy, as it provides a "fingerprint" of the protein showing the N-H bond in the backbone amide of each residue. Other two-dimensional experiments and isotope enrichment labeling schemes allow the methyl groups of particular residues to be used as probes. Methyl groups have favorable qualities for spectroscopy of larger protein complexes. Each methyl group has three protons, as opposed to one in an amide. Because methyl groups can rotate around an axis of symmetry, they have faster correlation times than the overall correlation time of the protein. This produces narrow linewidths even in large proteins (Boswell and Latham, 2018).

To date, few NMR studies of protein-polymer conjugates have been reported. Four PEGylated proteins have been studied by NMR: interferon α_{2b}

(Wang et al., 2000), human growth hormone-releasing factor (Digilio et al., 2003), interferon α_{2a} (Dhalluin et al., 2005) and plastocyanin (Cattani et al., 2015). Interestingly, each of these PEGylated proteins is modified using a different conjugation technique, though the strategy for characterization was very similar. First, the locations of the modifications were determined from using chemical shifter perturbations (CSPs), prior knowledge of free reactive groups or predefined chemistry. It is important to note, unlike the grafting-from ATRP-based PBPE, these conjugates only contained a single PEG moiety, and not multiple points of attachment. Thus, identification of the modified site is significantly simpler. Still, identification of positional isomers was the prime objective of two of the studies (Wang et al., 2000; Dhalluin et al., 2005). It was found that the PEG group acts independently from the protein (Cattani et al., 2015), and that positional isomers tended not to greatly affect the physical properties of mono-PEGylated proteins (Dhalluin et al., 2005). A brief overview of previously used experiments, as well as those relevant to this current study is presented in the following sections.

Chemical Shift Perturbations

CSPs are a simple, yet powerful tool in NMR spectroscopy. A CSP is defined as a change in the position of a resonance line in an NMR spectrum as a result of some treatment (ligand binding, pH, chemical modification). CSPs are often used in ligand-binding experiments to map the binding site and affinity (Williamson, 2013; Lisi and Loria, 2016). For the study of protein polymer conjugates, CSPs report on changes to the environment around each backbone amide, or side chain methyl. Chemical shifts are exquisitely sensitive to changes in this environment, which reflect changes in electron shielding. CSPs can be detected even if a binding site or structural perturbation is distal to the probe, reflecting allosteric effects. Chemical shifts can also be impacted by chemical

exchange (Section 3.2) if each of the states have unique chemical shifts. If the process is in slow exchange regime two resonance lines will be observed, but if the process is in the fast exchange regime a single line will be observed at the average frequency between the two states. This means if a peak shows a CSP in a protein-polymer conjugate it could be for two reasons: a change in the chemical composition (covalent modification) or a change in the environment (ligand/polymer binding). In studies of PEGylated proteins by NMR (Wang et al., 2000; Digilio et al., 2003; Dhalluin et al., 2005; Cattani et al., 2015), the authors report that CSPs are local to the site of PEG conjugation. This is quite interesting, as it suggests that the PEG does not bind to the surface of the protein. Previous studies of post-translationally modified proteins by NMR have shown that chemical modifications can be followed by NMR and provide quantitative information about site specificity and degree of modification (Smet-Nocca et al., 2010; Theillet et al., 2012; Kamah et al., 2014; Lee et al., 2016). The acylation of lysine residues, which is similar to attachment of the initiator using the NHS-ester reagent (see Section 1.1), results in a CSP to the backbone amide as well as the appearance of a lysine Nζ-acyl peak in the lower right of the amide region in the ^{15}N HSQC.

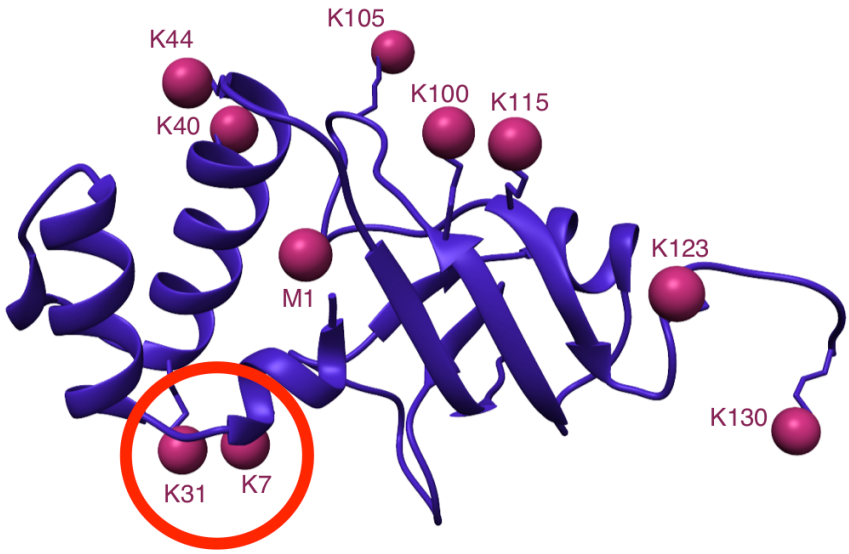
A prerequisite to obtaining CSPs is the assignment of resonances. These are usually obtained by a combination of NMR experiments which exploit connectivity (HNCA, HNCACB, TOCSY) and spatial (NOESY) effects to provide sequence specific assignments. One class of experiments, triple-resonance experiments, which provide residue identity and connectivity information exploiting characteristic carbon shifts for particular amino acids (Kay et al., 1990; Farmer et al., 1992). These experiments are named for the atoms involved in the magnetization transfer pathways. The HNCA experiment is a three dimensional experiment where the ^{13}C the Cα for a residue and the preceding residue is correlated to the ^{15}N and ^{1}H of the amide. This allows for adjacent residues

to be identified (e.g. Gly Cα appears at around 40 ppm in ^{13}C and Thr Cα at around 60 ppm) and by comparison to the sequence, contiguous portions of the protein can be assigned. The similar chemical shifts of Cα atoms are often ambiguous, making it difficult to obtain unambiguous amide N-H assignments, thus the HNCA is usually performed with other experiments such as the HNCO or HNCACB (Grzesiek and Bax, 1992). In the case of the HNCACB, another experiment used in this study, similar information to the HNCA is obtained, but rather than detecting Cα correlations, Cβ correlations are detected.

1.3 Rho130 as a Model System

The *E. coli* Rho factor (419 residues) forms a hexamer and contains an ATPase domain with structural homology as the F1 ATPase (Allison et al., 1998). Full-length Rho is an ATP-dependent DNA-RNA helicase required for RNA transcription termination. Rho130 is the first 130 residues of Rho, containing the RNA-binding domain of Rho. Rho130 was chosen as a model system because of its small size (14.6 kDa), stability during lengthy NMR experiments and the availability of full ^{1}H, ^{15}N and ^{13}C resonance assignments. Rho130 has been previously crystallized in both unliganded (Allison et al., 1998) and bound to rC3 (an RNA 3-mer of C) (Bogden et al., 1999) and has had NMR solution structure determined (Briercheck et al., 1998) (Figure 1.3A). The wealth of structural and spectroscopic information provides a benchmark to which polymer-conjugated Rho130 can be compared. The high density of positively charged residues (9 lysines and 9 arginines) allows for Rho130-initiator conjugates (Figure 1.3B) to maintain stability in solution, despite losing charges on modified lysines.

The objective of this study was to determine if charged polymers have an impact on protein structure and dynamics. The charged polymers that




```
         10         20         30         40
MNLTELKNTP VSELITLGEN MGLENLARMR KQDIIFAILK
         50         60         70         80
QHAKSGEDIF GDGVLEILQD GFGFLRSADS SYLAGPDDIY
         90        100        110        120
VSPSQIRRFN LRTGDTISGK IRPPKEGERY FALLKVNEVN
        130
FDKPENARNK
```

Figure 1.3: Unliganded NMR structure and sequence of Rho130. The structure of Rho130 is shown with the N-terminus and lysine Nζ (points of polymer attachment) represented in magenta spheres. A circle is drawn around K7 and K31 as they engage in H-bonds in many of the low-energy structures. This means initiator attachment at these sites may be hindered. The sequence of Rho130 is shown below with points of attachment highlighted in red.

were chosen for this study were short, carboxylate functionalized methacrylates (CAm) polymers (designated pCAm) of length 5-10 monomers. By comparing chemical shifts to that of unmodified proteins it was found that modification does alter the environment of nearly all residues in the protein. By examining both Rho130-initiator and Rho130-polymer conjugates, it was determined that the initial modification with the initiator group was showed the greatest CSPs, though the dynamics seemed to be more affected by the polymers.

Chapter 2

Characterization of Rho130 Conjugates

Rho130 is a "well-behaved" protein in solution, meaning it does not form aggregates and is soluble in water at high concentrations (>1 mM). This is due in part to its high number of charged residues, including the 9 lysines which are modified by the initiator. This makes Rho130 a good candidate for studying lysine-attached conjugates because it is able to maintain solubility despite modification. Although the charge at the protein surface is altered, Rho130-initiator conjugates are still well-behaved in solution. By synthesizing charged polymers from the surface of the Rho130-initiator conjugates, the surface of the protein is again altered, which does not seem to affect solubility greatly. Additionally, the modified lysines are distal to the site of RNA binding, possibly leaving its function intact (See Future Perspectives, Section 2.5). Though the protein seems to behave similarly, the ability to detect minute changes in protein structure and dynamics afforded by NMR provides a much more detailed view of the affect of the initiators and polymers. Significant effort was made to characterize the Rho130-initiator conjugates, as it is this step that is the

major determinant of heterogeneity for the conjugates. A number of methods for characterization of protein-initiator conjugates are discussed in Section 2.1.

CSPs showed that more prominent perturbations to the protein are observed after initiator modification than subsequent polymer growth. A discussion of the NMR spectra of Rho130-initiator conjugates as well as spectra along the initiator attachment reaction coordinate are presented in Section 2.2. Two Rho130-polymer conjugates will be discussed in the following sections: Rho130-pCAm and Rho130-pQA (Figure 2.1). Rho130-polymer conjugates were synthesized by Chad Cummings, Hironobu Murata and Bibifatima Kaupbayeva in the Russell laboratory at Carnegie Mellon University. More detailed studies of Rho130-pCAm are presented, as the synthesis of Rho130-pQA conjugates proved troublesome. Rho130-initiator and Rho130-polymer conjugates maintain their structure as shown by a well dispersed ^{15}N HSQC and ^{13}C HMQC (Sections 2.2 and 2.3). However, widespread chemical shift perturbations suggest possible structural or dynamical changes are brought about from modification of Rho130. Finally, Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion was used to probe possible perturbations to μs-ms dynamics (Section 2.4).

2.1 Determining Degree of Modification

Initiator attachment, the first step in PBPE, involves the reaction of an NHS ester modified form of an ATRP initiator with the primary amines of a protein. Rho130 contains 10 primary amines (9 lysine residues and the N-terminus) as points of attachment. Rapid determination of degree of modification was critical in developing reaction conditions which produce uniformly labeled protein. Because the polymers only extend from the initiator, it is important to minimize positional isomers by ensuring near complete coverage of the primary

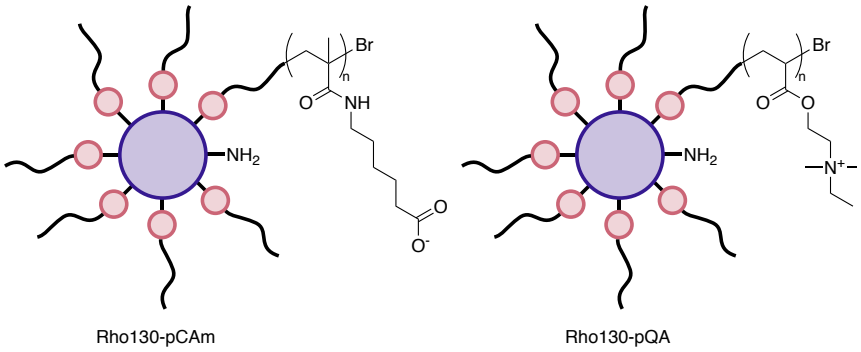


Figure 2.1: Schematic View of Rho130-polymer Conjugates. Rho130-pCAm (left) and Rho130-pQA (right) conjugates are represented with the protein in purple, initiator group in red and polymer as black curves. The absence of an initiator group on one of the primary amines reflects the fact that the lysine side chains are not 100% modified, leading to positional isomers of the Rho130-polymer conjugates.

amines. Two methods were used to determine the number of initiators attached to Rho130: electrospray ionization mass spectrometry (ESI MS) and polyacrylamide gel electrophoresis (PAGE).

Development of Assays for Degree of Modification

ESI MS uses a "soft ionization" method so as not to fragment the analyte, leaving proteins intact and allowing for molecular weight determination of full-length proteins. Because the mass to charge ratio (m/z) is reported and proteins contain many ionizable groups, proteins in the kDa-MDa range are routinely studied using this technique (Pitt, 2009). The mass spectrum of unmodified Rho130 is shown in Figure 2.2. As expected, after initiator attachment the charge envelope is shifted to higher (m/z) as the addition of the initiator both add weight and removes the positive charge from the lysine side chain. From these spectra it was determined that the major products of the initiator attachment reaction are conjugates with 8 or 9 initiators attached, though the distribution of products varied greatly from batch to batch.

While the most precise method, ESI MS proved labor intensive when testing multiple reaction conditions requiring a simpler, more efficient method for determine degree of modification. By performing PAGE in non-denaturing and non-reducing conditions, proteins are left intact and migrate according to both their size and charge. Because a positive charge is removed upon addition of the initiator, each conjugate migrates with a distance based on the number of initiators. By confirming a "standard" Rho130-initiator conjugate sample by ESI MS, the number and distribution of initiators could be quickly determined using PAGE. As a proof of concept the gel in Figure 2.3 shows a time course of the initiator attachment reaction. The number of initiators and distribution are

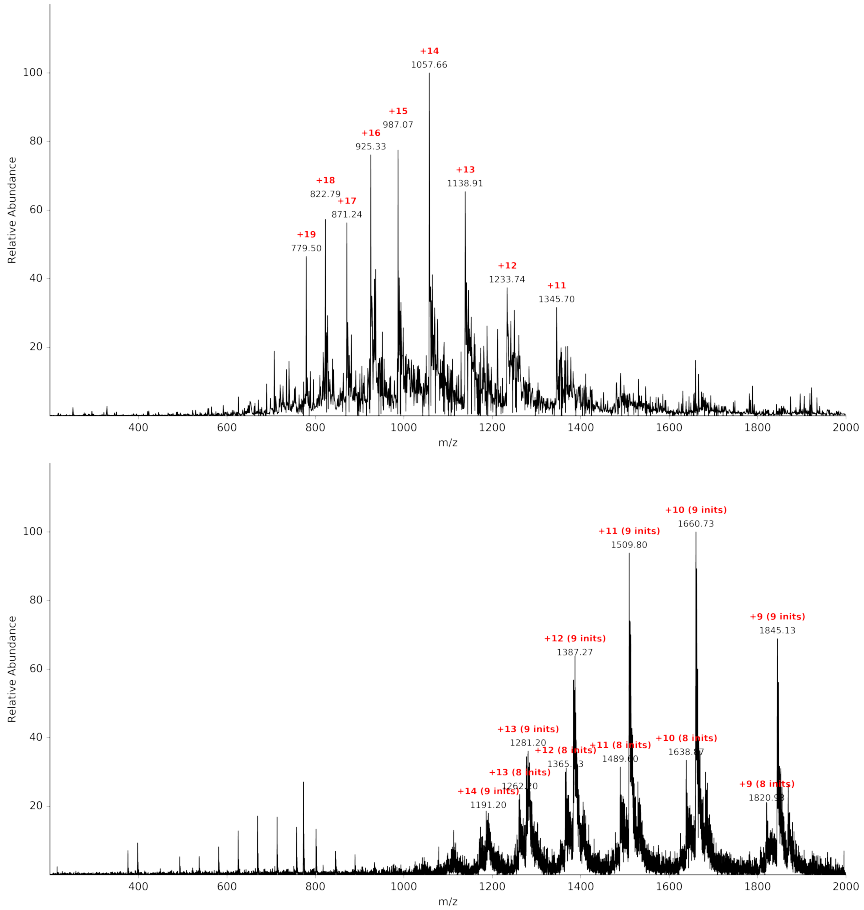


Figure 2.2: Unmodified Rho130 and Rho130-initiator conjugate ESI mass spectrum. The ESI spectrum of ^{15}N-labeled Rho130 (top) and Rho130-initiator conjugates (bottom) with charge envelope labeled in red. In the initiator spectrum, multiple conjugates can by identified. The movement toward higher m/z reflects an increase in molecular mass, and more importantly loss of positive charge from modification of the lysine residues. The envelope of unmodified Rho130 contains much higher charged states.

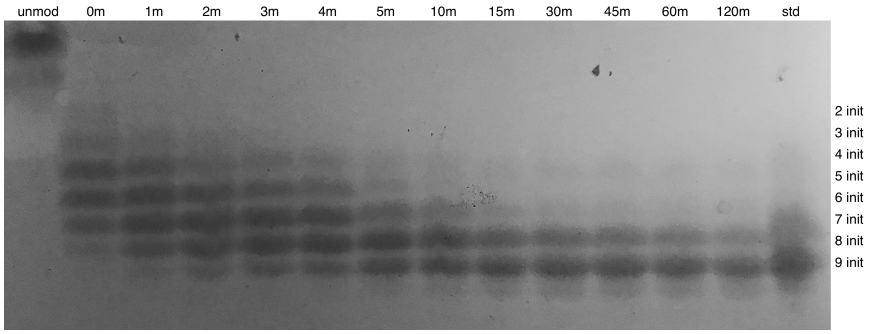


Figure 2.3: Time course of initiator attachment visualized by PAGE. A time course reaction, quenched at each timepoint by addition of free lysine is shown in each lane. Using PAGE, the reaction can be followed, or multiple reaction conditions can be assayed at once. Each band on the gel represents the addition of 1 initiator molecule. Rho130 in the unmodified lane contains an impurity, making assignment of the 1-initiator band uncertain.

easily visualized at each time point, or for a PAGE gel with lanes corresponding to different reaction conditions: the major products of the reaction.

Optimization of Initiator Attachment

Using these techniques to assay degree of modification an optimized protocol for initiator attachment was developed. A series of conditions for the reaction were tested with a 50-fold excess of NHS-initiator reagent as a DMF solution (10% v/v in the final buffer) to Rho130 at pH 7.8 producing the narrowest distribution of products. The products of this reaction are a 3:1 mixture of Rho130-initiator conjugates with 9 or 8 initiators attached, as determined from ESI MS (Section 2.1). Subsequent growth of polymers from these Rho130-initiator conjugates resulted in well-resolved ^{15}N HSQC spectra as opposed to those produced before optimization of the initiator attachment step (Section 2.3). This provided a suitable sample for further structural and dynamical studies of Rho130-polymer conjugates. Though the number of initiators per protein molecule could be determined using the facile PAGE method, the location and percent modification could not be easily or quickly determined.

Initiator Mapping by Proteolysis and Mass Spectrometry

Mapping sites of initiator attachment is critical to predicting which sites polymers will be grown from and how this may effect the protein's behavior. In proteomic studies of post-translationally modified proteins, MS is integral to identification of modified residues and which of the many post-translational modifications occur at each of these sites (Han et al., 2008). Typically tandem MS coupled with liquid chromatography is employed on trypsin-generated peptides to provide maximum coverage over the proteome. Fortunately, when using a single protein with a single type of modification, the process is considerably simpler. Though, trypsin digestion of Rho130 did not produce sufficient mass

spectra likely due to the large number of lysines and arginines resulting in very small cleavage products. For this reason LysC, a protease which cleaves C-terminal to lysine residues, was used to generate the peptides (Wada and Kadoya, 2003). For unmodified Rho130 3 distinct peptides can be detected by ESI MS: N8-K31 (cleavage at K7 and K31), Q32-K40 (cleavage at K31 and K40), S45-K100 (cleavage at K44 and K100), providing coverage over five of the lysines: K7, K31, K40, K44 and K100 in the control (Figure 2.4A). This control experiment provides evidence that Rho130 can be digested by LysC, and the same protocol was used for Rho130-initiator conjugates. In the Rho130-initiator conjugates cleavage at a lysine suggests the lysine is unmodified, as the initiator group would sterically hinder the protease. The ESI MS of Rho130-initiator conjugate LysC peptides is shown in Figure 2.4B, and resembled the mass spectrum of un-cleaved Rho130-initiator conjugates (Figure 2.2B). Notably, the m/z is slightly different, corresponding to loss of the first 7 residues (cleavage at K7). This means that in the majority of Rho130-initiator conjugates are unmodified at K7. A second, much less abundant, peptide is also present in this spectrum corresponding to residues N8-K31 (cleavage at K7 and K31). This means that the minor 8-initiator product is unmodified at K7 and K31.

A second, complementary protease was chosen in an attempt to remedy the coverage issues imposed by LysC. This protease, GluC, cleaves C-terminal to glutamic acid residues, and to a lesser extent aspartic acid residues. This avoids the issue of steric hinderance causing bias in the data, and allows for unmodified residues to be observed directly rather than inferred. The GluC peptides from unmodified Rho130 offered coverage over K40, K44, K115 with 8 peaks in the mass spectrum. The mass spectrum of Rho130-initiator conjugates resulting in 12 peaks in the mass spectrum, and due to the large number of combinations between modified/unmodified lysines, cleavage at Glu or Asp,

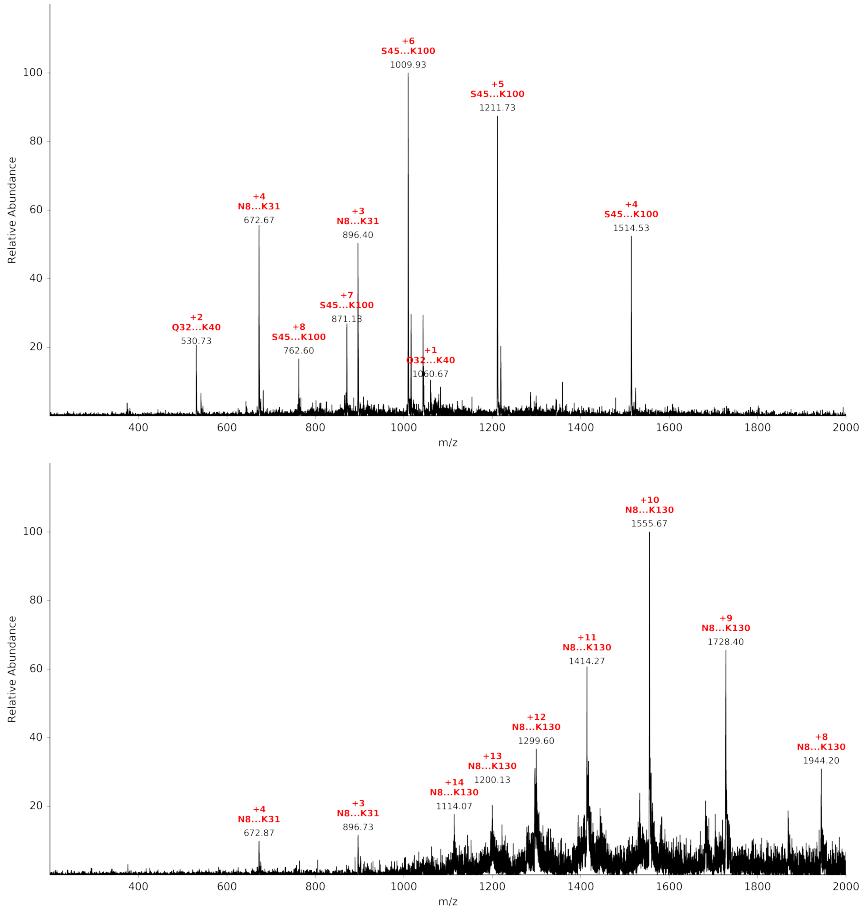


Figure 2.4: ESI MS of Unmodified and Initiator-Modified Rho130 after LysC Digestion. Unmodified Rho130 (top) is cleaved into at least 3 identifiable peptides by MS, giving coverage over K7, K31, K40, K44 and K100. Rho130-initiator conjugates (bottom) show a major cleavage product of N8-K130 (cleavage at K7), suggesting it is unmodified, and a minor cleavage product N8-K31 (cleavage at K7 and K31). Cleavage suggests the lysine is unmodified, as the initiator group would sterically hinder the LysC protease. This suggests that the 9-initiator conjugate is modified at all sites but K7 and the 8-initiator conjugate is modified at all sites but K7 and K31.

and missed cleavages, the results were not easily interpreted. Tandem MS was performed on these samples, though their analysis remained challenging and did not provide straightforward assignment of fragments. The quality of the data from this experiment were of decent quality and repetition with an automated method of peak assignment and LC MS-MS may provide the level of detail that was hoped to be gained.

2.2 NMR Spectra of Rho130-initiator Conjugates

A benefit to using Rho130 as a model system is that nearly complete assignments for each atom in the protein have been previously reported (Briercheck et al., 1996). This provides a benchmark to compare to upon modification of the protein, and the resolution to determine effects to nearly every residue. For this study two types of probes were chosen: backbone amides and the methyl groups of isoleucine, leucine and valine (ILV) residues. These provide information on both the protein backbone as well as the side chains, to give a more detailed view of possible alterations to structure and dynamics. Focusing on only the ILV methyls reduces the number of assignments to be made after initiator attachment, as very few peaks show CSPs. Additionally, ILV methyls serve as excellent probes in higher molecular weight complexes such as the Rho130-pCAm conjugate which are more than double the molecular weight of unmodified Rho130.

Assignment of Rho130-initiator conjugate ^{15}N HSQC

Immediately noticeable in the ^{15}N HSQC of Rho130-initiator conjugates (red spectrum Figure 2.5) are the presence of additional resonance lines at 7.9 ppm ^{1}H, 128 ppm ^{15}N. These peaks correspond to the amide bond formed between the initiator molecule and lysine Nζ. Fortuitously, this provides a

means of determining the degree of modification at each lysine, granted they can be assigned. Assignment of the initiator-lysine Nζ peaks was achieved by a mutagenesis, where a series of K$\rightarrow$R mutations were made, expressed in ^{15}N-containing media, subjected to the initiator attachment reaction and a ^{15}N HSQC performed on the products. The assignments are shown in Figure 2.6 and marked with an asterisk (*). Unfortunately, K44 and K100 overlap and peak intensities cannot be extracted for each. Also mutagenesis of K7 to arginine does not result in the loss of a peak in the spectrum supporting the mass spectrometry data after digestion with LysC (Section 2.1).

Also apparent from the Rho130-initiator conjugate amide spectrum are the widespread CSPs (Figure 2.6). Most peaks in the spectrum have moved, though most are fairly minor perturbations. In order to match shifted peaks with their original positions, standard triple-resonance experiments (HNCA and HNCACB) were used to assign the Rho130-initiator amide peaks. This also allows for the assignment of the N-terminal methionine to be determined, as it becomes an amide after reacting with the initiator. Assignment of nearly all peaks in the spectrum was achieved by comparison to the unmodified spectrum and triple-resonance experiments. Interestingly, upon modification with the amides of I49, V116, and N117 are exchange broadened and no longer visible in the ^{15}N HSQC. All of these residues are proximal to modified lysines, suggesting some change in the backbone dynamics. Especially in the vicinity of K115, which is flanked by two exchange broadened residues.

Structural and Dynamical Changes to Rho130-initiator Conjugates

Major CSPs are observed for the amides N2, L3, Q41, H42 and K44. CSPs near the N-terminus are understandable, as it is modified on the backbone and not the side chain. This may lead to structural perturbations for nearby residues

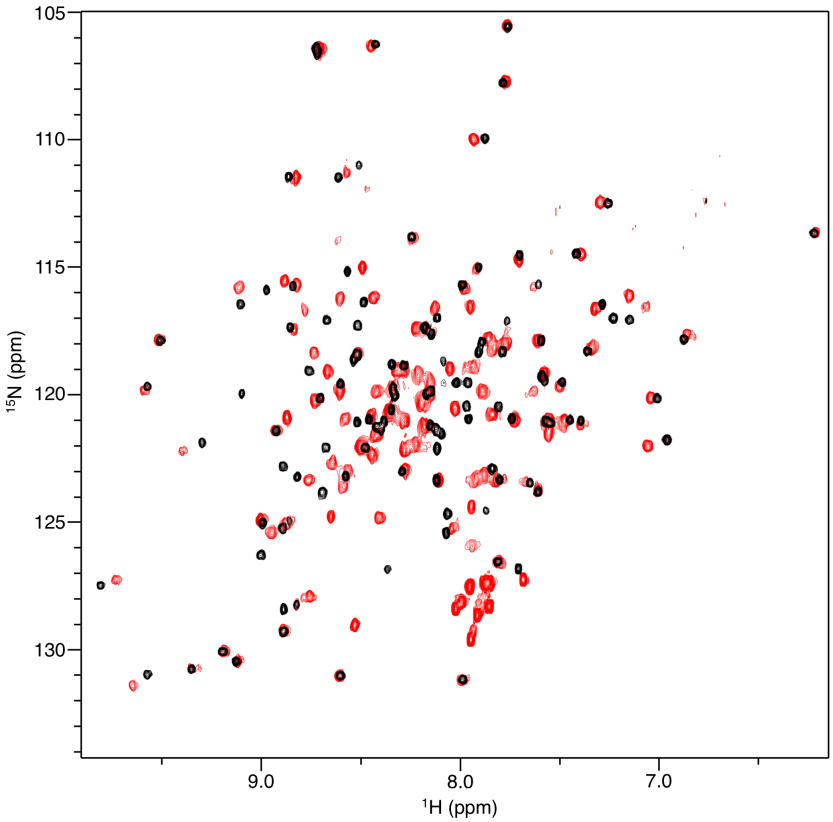


Figure 2.5: ^{15}N HSQC of Rho130-initiator conjugates. The unmodified Rho130 spectrum is shown in black and Rho130-initiator conjugate in red. Significant CSPs are observed. Labels are omitted here for clarity and a spectrum with assignments is shown in Figure 2.6.

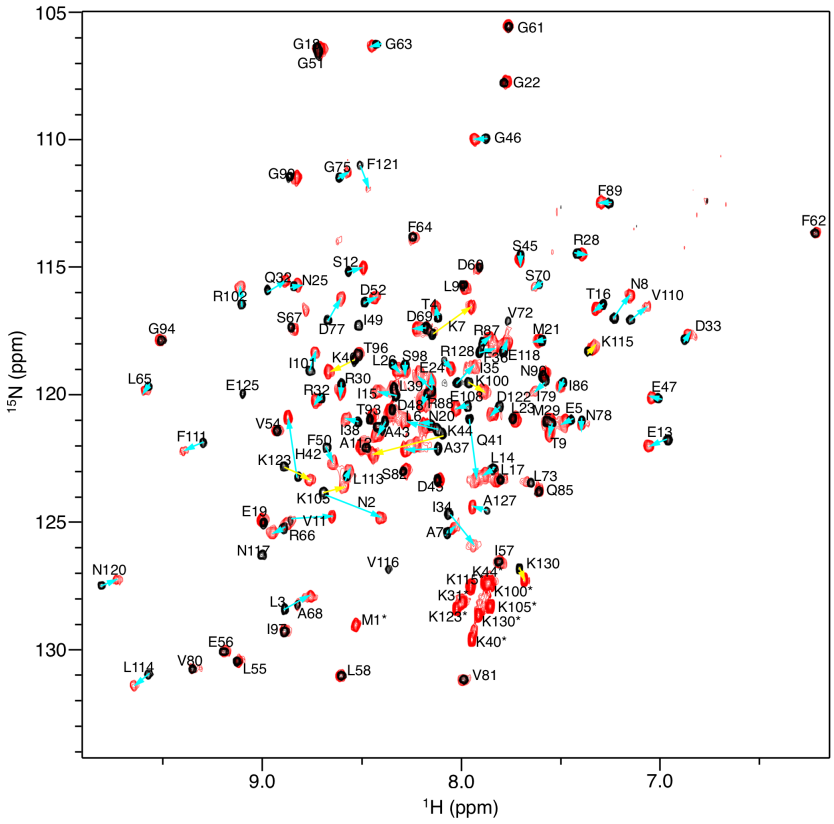


Figure 2.6: ^{15}N HSQC of Rho130-initiator conjugates. The unmodified Rho130 spectrum is shown in black and Rho130-initiator conjugate in red. Chemical shift perturbations are shown as cyan arrows for most residues. Yellow arrows are used to highlight CSPs for lysine backbone amides. Widespread CSPs reflect the coverage of the attachment of the initiator. Major CSPs are generally local to the points of attachment. The spectra of the unmodified protein and initiator conjugate were made from the same stock of purified Rho130.

to accommodate the bulky initiator group. Large CSPs are also expected at positions 40-44 as both K40 and K44 are potential sites of modification. These residues are situated at the end of α3, possibly indicating local unfolding of the helix induced by the chemical modification.

By comparing the volumes of the initiator-lysine Nζ peaks, the site specific degree of modification can be determined (Figure 2.7). Because K7 does not appear in the spectrum, and digestion with LysC followed by mass spectrometry shows the major species in solution is the protein cleaved at K7, it is assumed to be negligibly modified. As the products of the initiator attachment reaction are a mixture of 3:1 9- and 8-initiator conjugates, it is expected that the degree of modification would be nearly 100% for each lysine. Interestingly, and again consistent with the LysC data, K31 has a peak intensity of half of K130 meaning that it is usually the unmodified residue in the 8-initiator conjugate. K40 and K44 are very close to one another in the 3D structure, suggesting steric hinderance may play a role in reducing its chance of being modified. This is reflected in a reduced peak intensity as compared to the other lysines. Unfortunately, because of the overlap between K44 and K100 it is impossible to determine their degree of modification using this method. It is likely that local dynamical differences result in different peak intensities for each residue and if the major product of the reaction is a 9-initiator conjugate all of lysines are modified with the exception of K7. K7 is likely less reactive because of its interaction with a nearby loop. Interestingly, the backbone amide of K7 does show a significant CSP, suggesting a structural perturbation at this location that is not a consequence of K7 modification. While perhaps not entirely accurate, quantification of peak intensities from the ^{15}N HSQC does serve as a measurement of relative reactivity for each of the lysines. For the most part the solvent accessible area and degree of modification correlate quite well (Figure 2.7). The fact that K7 is unmodified and K31 and K40

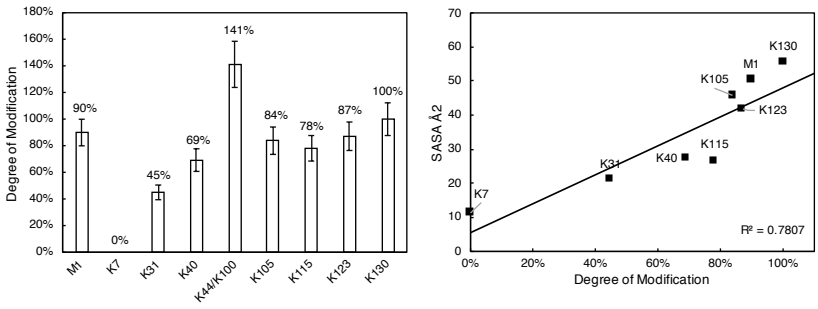


Figure 2.7: Degree of Modification with Initiator Determined by NMR. Percentage modification is determined by normalizing lysine side chain peaks to the side chain peak of K130 (hence 100% at that position). M1 was normalized to a nearby isolated backbone peak (L58) as the intensity of the side chain peaks is an order of magnitude higher. An error of 10% is shown on each bar, estimated by normalizing the intensity of the side chain peak of K130 to its backbone peak, which is 10% less intense. K44 and K100 are overlapped in the spectrum and individual intensities could not be extracted for both. Solvent accessible area and degree of modification are well correlated.

are less modified is consistent with the solvent accessible area for each primary amine. Though K115 also has a comparable solvent accessible area, and appear to have no issue with reactivity.

ILV Methyl Spectra of Rho130-initiator Conjugates

ILV methyl CH_3-labeled Rho130-initiator conjugates were also produced to observe any changes to the hydrophobic side chains of Rho130 upon attachment of the initiator. These methyl groups serve as ideal probes of residues often found in the hydrophobic core of the protein, and provide complementary information to the backbone amides. Very few CSPs are observed in the Rho130-initiator conjugate spectrum (red spectrum Figure 2.8), thus peaks were assigned by comparison to the unmodified spectrum and required no additional experiments. As expected, residues near modified primary amines show CSPs (L3, I49, I97, L113, L114, and V116). Notably ,V11 and L17, which reside on α2 which makes contact with α1 near K7 and α3 near K31, also show CSPs suggesting dynamical changes to the helical portion of Rho130 (residues 1-45). It is possible that the normal behavior of these helices is impacted by conjugation at the N-terminus and K31. This may also reflect the heterogeneity in K31 modification, indicated by the broadening of the L17 peak. I49 and V54 also show significant line broadening, indicating changes to dynamics in the first sheet of the β region (residues 50-54). Interestingly, I49 is exchange broadened in the ^{15}N HSQC and V54 shows some broadening, though it is indistinguishable from the global line broadening observed for all peaks in the spectrum.

Monitoring Initiator Attachment Kinetics by NMR

Specific labeling of lysine residues can be accomplished by growing *E. coli* with a ^{14}N nitrogen source and adding [^{13}C, ^{15}N] lysine. The absence of ^{15}N scram-

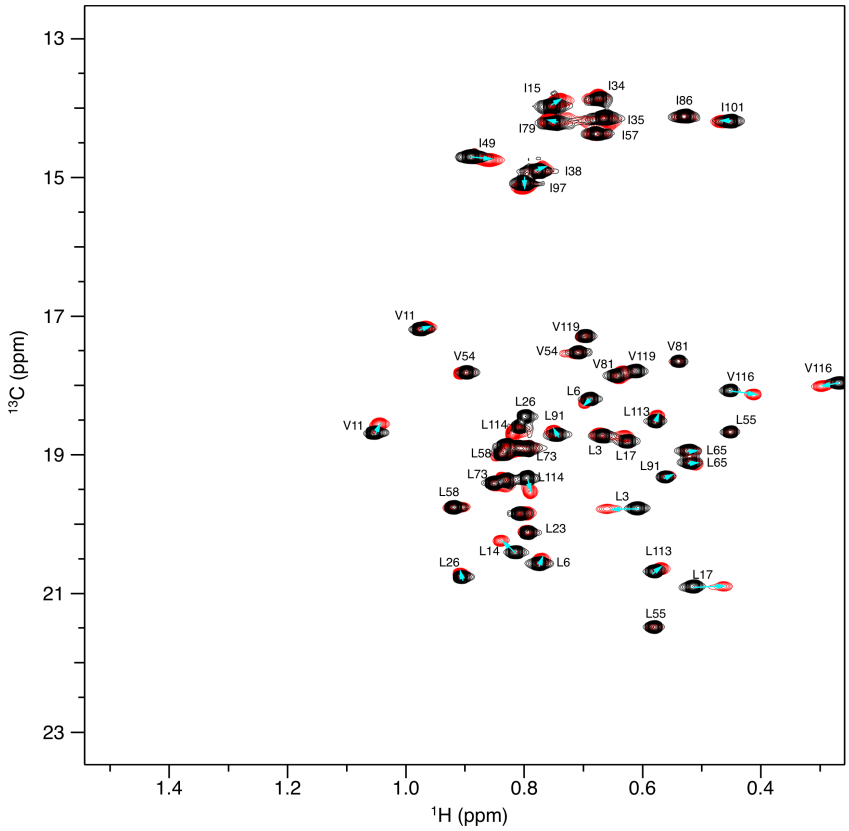


Figure 2.8: ^{13}C HMQC of Rho130-initiator conjugates. The unmodified Rho130 spectrum is shown in black and Rho130-initiator conjugate in red. Chemical shift perturbations are shown as cyan arrows. The side chain methyl CSPs are mostly localized to residues near initiator-modified lysines. Though some (V11, L17) seem to indicate larger dynamical changes in the helical portion of Rho130 (residues 1-45). Also apparent is the broadened lines in the Rho130-initiator conjugate spectrum. This suggests changes to the slow dynamics or heterogeneity.

bling means that only lysines are visible in the ^{15}N HSQC. This significantly reduces the complexity of the spectrum and allows both the backbone and side chain lysine peaks to be monitored throughout the course of the reaction. Because the reaction of an NHS-ester with a primary amine on a protein occurs quite rapidly, samples were prepared by quenching the reaction at multiple time points with an excess of free lysine. The same 5x molar excess of initiator (per lysine) was used in this reaction, which quickly went to completion. Interestingly, the backbone peaks for K7, K44, K100 and K130 undergo a two-part CSP suggesting the final position seen in the fully modified protein may be the result of several structural intermediates. It is entirely possible that more positions on the trajectory can be filled in using sub-stoichiometric quantities of the initiator reagent, which would provide site specific reactivities for each lysine. Interestingly, K7 which is supposedly unmodified does undergo a significant CSP, as observed in the full ^{15}N HSQC and seems to exist in two states in the fully modified Rho130. Though the backbone amide of K31 is not observable in either spectrum, the two K7 backbone peaks may represent the nearly 50/50 modification of K31 as they are near one another in the 3D structure.

2.3 NMR Spectra of Rho130-polymer Conjugates

Fortunately, very few CSPs are observed after performing ATRP on the Rho130-initiator conjugates (Figure 2.11). This suggests two things: (1) the structure of the protein differs very little between these two conjugates and (2) attachment of the initiator is more disruptive than subsequent polymerization. This is notable, as it suggest that if proteins are amenable to modification with the ATRP-initiator, then their structure should be well preserved after the polymers are grown. The ^{15}N HSQC of Rho130-pCAm conjugates shows broadening of most resonances, while CSPs seem to be limited to modified residues

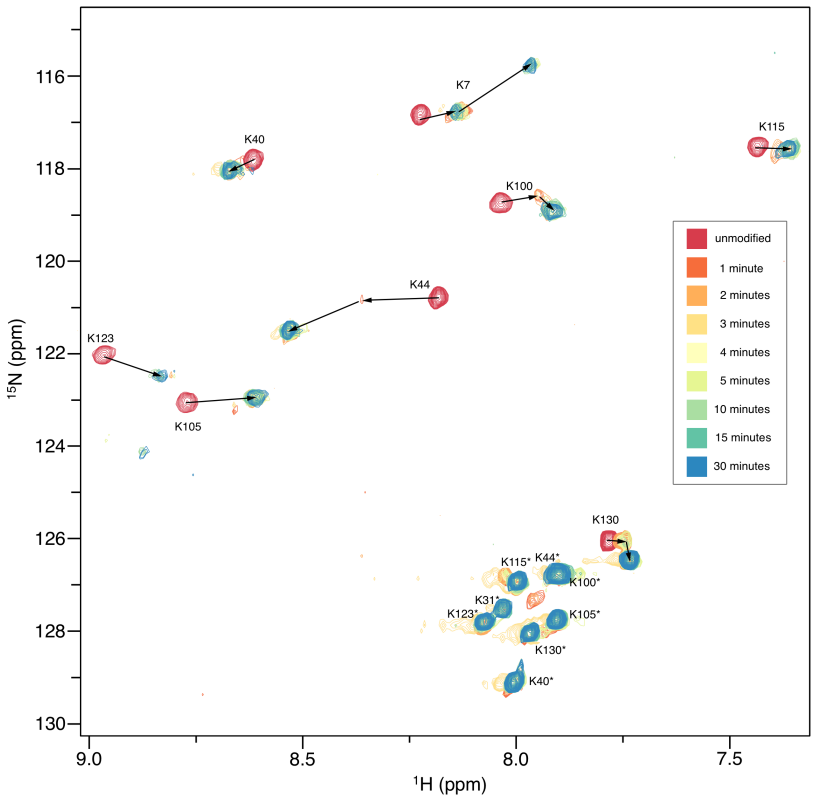


Figure 2.9: Lysine ^{15}N HSQCs of Initiator Attachment Time Course. Specifically lysine labled Rho130 was subjected to the initiator reaction, quenched at various timepoints (0, 1, 2, 3, 4, 5, 10, 15 and 30 minutes). ^{15}N HSQC spectra were collected on each quenched sample. Notably, a two-part CSP is seen for the backbone resonances of K7, K44, K100 and K130. This may indicate multiple structural rearrangements after modification.

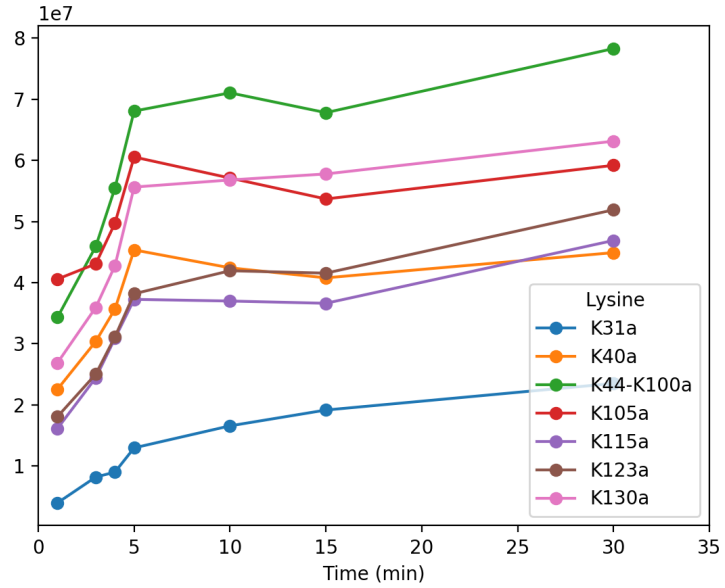


Figure 2.10: Lysine Nζ-initiator peak intensities over time course. The peak intensities from the spectrum in Figure 2.9 are plotted as a function of time. Though not fit to a kinetic model, the rate of K31 modification is clearly slower than the rest. This reflects the local structural effects that protect it from reacting with the initiator reagent.

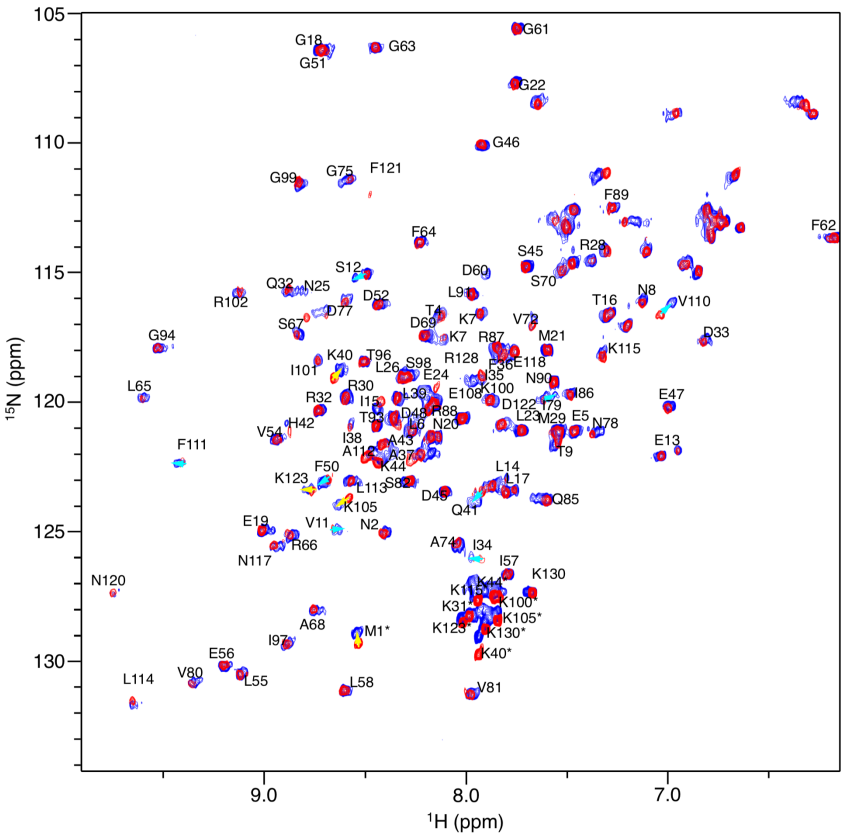


Figure 2.11: ^{15}N HSQC of Rho130-pCAm conjugates. The Rho130-initiator conjugate spectrum is shown in red and Rho130-pCAm conjugate is in blue. Chemical shift perturbations are shown as cyan arrows for most residues. Yellow arrows are used to highlight CSPs for lysine backbone amides and M1. There are very few CSPs between the initiator and polymer conjugates, though most of the amides are linebroadened in the Rho130-pCAm spectra. Additional peaks in the top left corner (in comparison to Figure 2.6) are from side chain amides and are not assigned. The spectra of the both conjugates was performed on the same protein at each step of the reaction.

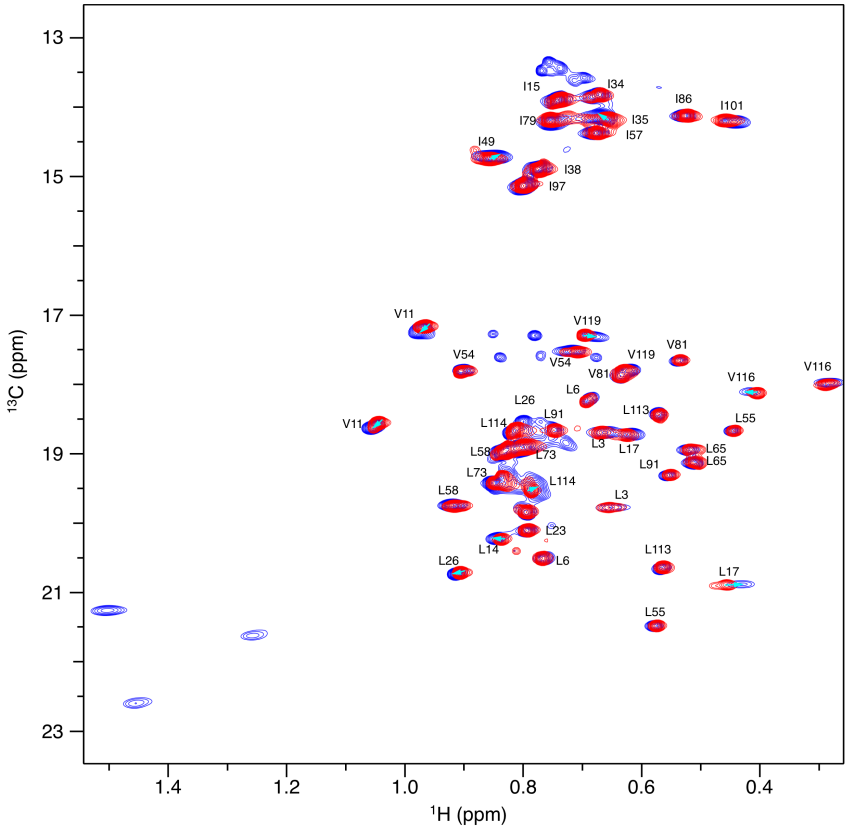


Figure 2.12: ^{13}C HMQC of Rho130-pCAm conjugates. The Rho130-initiator conjugate spectrum is shown in red and Rho130-pCAm conjugate spectrum in blue. Chemical shift perturbations are shown as cyan arrows. Similar to the ^{15}N HSQC, fewer perturbations are observed from the initiator to polymer-modified spectra than from the unmodified to initiator-modified spectrum. The presence of line broadening and spurious resonance lines near the middle of the isoleucine and leucine/valine regions indicates heterogeneity. The peaks from 1.2-1.5 ppm ^{1}H are likely natural abundance peaks from the methyl group of the CAm polymer. Interestingly, these peaks show chemical shift dispersion rather than appearing as a single peak, suggesting polymers in multiple different environments.

and those that are proximal to the sites of polymer growth. Also, the region of the spectrum corresponding to the polymer-lysine Nζ amides is significantly broadened and heterogeneous. As expected, the differing lengths of polymers produced by ATRP may affect the lysine side chains in different ways, leading to multiple states being observed for each. The ILV methyl spectrum tells a similar story. However, it is clear from the spurious peaks around 13 and 17 ppm ^{13}C, that many different states of the ILV side chains are present, though assignments for these peaks have not been determined. The additional peaks from 1.2-1.5 ppm ^{1}H correspond to natural abundance ^{13}C correlations of the methyls present on the backbone of the pCAm monomers (Figure 2.1). The chemical shift dispersion for these peaks suggests the polymer backbone may exist in multiple environments, influenced by its location on the protein. The increase in molecular weight (approximately 32 kDa overall, compared to 14.6 kDa for Rho130) predicted by the polymerization conditions used to produce these conjugates results in line broadening for most of the peaks in the spectrum. Interestingly, as in the Rho130-initiator conjugate ILV methyl spectra, V11 and L17 show CSPs, indicating that the addition of polymer at K31 may further alter the dynamics of α2.

Optimization of Initiator Attachment Increases Spectral Quality

As discussed in Section 2.3, considerable effort was put toward optimization of initiator attachment. This was informed by the similarities between the Rho130-initiator and Rho130-pCAm conjugates NMR spectra, and the requirement of mostly uniform conjugates for further studies by NMR. Rho130-polymer conjugates synthesized using heterogenous mixtures of Rho130-initiator conjugates produce spectra such as those in Figure 2.13. While the backbone amides are all in the correct locations, many of them are broadened or split,

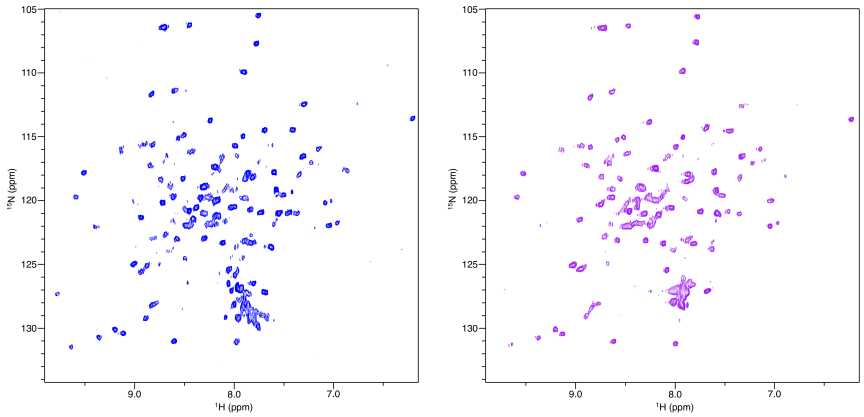


Figure 2.13: Rho130-polymer Conjugates before Optimization of the Attachment Reaction. The Rho130-pCAm (blue) and Rho130-pQA (purple) conjugates synthesized from Rho130-initiator conjugates before development of the optimized protocol produce poor quality NMR spectra.

leading to very poor quality spectra. The lysine side chain peaks also show quite a bit of variability as opposed to those from a more carefully controlled reaction (Figure 2.11). Optimization of initiator attachment, while perhaps not as exciting as introducing new functionalization through polymerization, is a crucial step toward understanding the behavior of these complex materials in solution. By taking the time to optimize the first step of PBPE and characterize the Rho130-initiator conjugate, Rho130-polymer conjugates suitable for further study could be produced. While the relaxation measurements described in the next section focus on Rho130-pCAm conjugates, it possible to modify Rho130 with any number or combination of polymers and study their behaviors.

2.4 CPMG Relaxation Dispersion on Rho130, Rho130-initiator and Rho130-pCAm Conjugates

Differences in the μs-ms dynamics of unmodified Rho130, Rho130-initiator conjugates and Rho130-pCAm conjugates were examined using CH_3 methyl CPMG relaxation dispersion. The CPMG experiment (further described in Section 3.2) is a windowed experiment, meaning that motions outside the range of approximately 100 to 5000 s^{-1} are not detectable by this method (Kleckner and Foster, 2011). This limitation is placed on the CPMG because R_{ex} is quenched when the pulse rate (ν_{CPMG}) is at approximately the same as or a faster rate than the exchange process. Few residues in Rho130 or the two conjugates show any exchange on these time scales. Fits were performed using `frenchdip` (Chapter 5) and show very little difference in reduced χ^2 between an exchange model and the "No Exchange" model. This was unexpected as the ^{13}C HMQC indicates changes to the dynamics of both the Rho130-initiator and Rho130-pCAm conjugate in line broadening in particular peaks (i.e. L17).

That is not to say these residues are not behaving differently, however they are outside the detectable limit imposed by the CPMG time window.

An exception to this is the residue I97, which can be fit to the fast exchange model. The fitted exchange parameters for I97 for unmodified Rho130, Rho130-initiator conjugates and Rho130-pCAm conjugates are summarized in Table 2.1. From the dispersion curves (Figure 2.14) and fitted parameters it is apparent that the Rho130-initiator conjugate at unmodified Rho130 behave similarly with a k_{ex} of approximately 3000 s^{-1}. Interestingly, I97 in the Rho130-pCAm conjugate undergoes a much slower exchange process at 1000 $\pm$ 300 s^{-1}. This suggests the polymer does, in fact, have some influence on the dynamics of the protein. The presence of 3 polymerization sites proximal to I97 may suggest why dynamical perturbations can be measured at this residue. Measurement of degree of modification at these three sites, K100, K115 and K123 (Figure 2.7), suggest they are fairly uniformly modified with initiator, as opposed to others like K31 and K40. The uniformity of the conjugate in this region may allow for dispersion to be measured for a single species rather than a more complex mixture. Alternatively, it may just be the rate of the motion at this position is within the time scales accessible to the CPMG experiment.

2.5 Conclusion and Future Perspectives

The remarkable ability for proteins to maintain their structure even after heavy chemical modification is a testament to the power of the cooperative forces that dictate protein folding. In the case of Rho130, a protein fragment, these forces are able to withstand having nearly every primary amine functionalized with a poly-anionic group with minimal changes to the overall structure and dynamics. Optimizations of the two-step ATRP-based PBPE workflow allowed for production of sufficiently uniform conjugates which could be used in NMR experiments that are incredibly sensitive to small populations of the minor state.

Protein	k_{ex} (s^{-1})	ϕ_{ex} (s^{-2})	R_2^0 (s^{-1})	χ^2/N
Rho130	3800 ± 1600	5300 ± 4900	7.0 ± 0.33	10
Rho130-Initiator	2900 ± 780	5800 ± 1800	8.1 ± 0.18	8.2
Rho130-pCAm	1000 ± 300	1600 ± 520	11 ± 0.10	6.5

Table 2.1: Exchange Parameters for I97

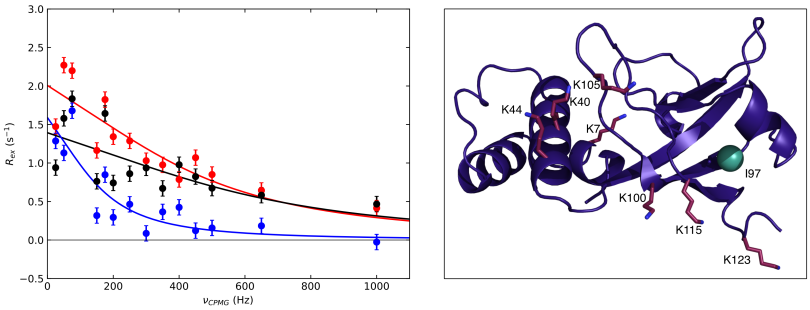


Figure 2.14: I97 shows a unique RD profile in Rho130-pCAm conjugates. In contrast to other residues, I97 is the only methyl group which produces a CPMG dispersion curve. The unmodified (black curve) and initiator conjugate (red curve) are more similar and can be fit to similar exchange parameters. The Rho130-pCAm conjugate (blue curve) has unique exchange parameters, suggesting the polymer is affecting the dynamics of the surrounding region. The methyl group of I97 is shown as a green sphere. Lysines are rendered as red sticks. The region surrounding I97 contains 3 potentially modified lysines, which may be responsible for the change in dynamics of the region in the polymer conjugate.

Through these experiments, insight was gained into the effect of extreme chemical modifications to the surface of the model protein Rho130 along each stage of the process. Initiator attachment seems to have a dominant effect on the structure of the protein, that is ATRP performed to Rho130-initiator conjugates results in little to no changes in both the ^{15}N HSQC or the ^{13}C HMQC. However, CPMG relaxation dispersion revealed that although they may be structurally similar to the initiator conjugates, the dynamics of Rho130-pCAm conjugates are clearly affected by the polymers grown from these initiators. This begs the question of whether this is a more widespread phenomenon. Protein-polymer conjugates have been gaining traction in the pharmaceutical and biotechnological spaces over the passed few decades. Losses in protein activity are often attributed to blockages of the active site, or other steric effects, but rarely are the dynamics of the protein considered.

In order to tease apart the remaining questions, a few experiments are proposed. The relatively minor CSPs suggest that the overall fold of Rho130 is essentially unaffected, with some possible unfolding upon attachment of the initiator around the C-terminal end of the α3, where two sites of modification are situated. An experiment which could provide some complementary information would be a methyl NOESY, which would provide structural information complementary to simple CSPs. A 3D amide NOESY was performed, however the intense side chain lysine peaks produce artifacts throughout the spectrum, greatly limiting its usefulness. Though the CPMG experiment only detected a single methyl group undergoing exchange on the μs-ms time scale, it is possible that other residues are, in fact, affected by modification at each step of PBPE. It is possible that at a different time scale, either faster or slower, dynamical differences between the unmodified, initiator and polymer conjugates could be more apparent. Also, the activity of Rho130 was not assayed after initiator or polymer modification. A titration of rC3 into the protein and measurement

of NMR peak intensities should provide a simple and robust method for determining the binding constant with both conjugates. Other open questions remain, as this study has just barely scratched the surface. Two, in particular, stand out among the rest: (1) what is the effect of polymer identity on protein dynamics? and (2) are the dynamics of the protein more perturbed by larger polymers? It is possible that future NMR studies are a reasonable route toward answering these questions.

Chapter 3

Human Thymidylate Kinase (hTMPK) Introduction

Thymidylate kinases (TMPKs, EC 2.7.4.9) catalyze the simple, yet essential phosphorylation of TMP to form TDP using ATP and Mg^{2+} (Reichard, 1988). While other deoxyribonucleotides can be produced from their counterpart ribonucleotides by ribonucleotide reductase, thymidine deoxyribonucleotides are formed exclusively by serial phosphorylation of TMP from one of two starting points: phosphorylation of thymidine by thymidine kinase using the salvage pathway or conversion of dUMP to TMP by thymidine synthase in the de novo pathway (Reichard, 1988; Jordan and Reichard, 1998). TMPK sits at the intersection of these two pathways serving as the only means of TDP production, and thus regulates the pool of available TTP for use in DNA synthesis and repair (Huang et al., 1994; Ke, 2005; Hu et al., 2019). All reported TMPKs are homodimers with an overall molecular weight of approximately 50 kDa, and are composed of four functionally important regions: the ATP α- and β-phosphate binding loop (P-loop), adenosine ring binding loop (A-loop), ligand induced degradation (LID) domain, and nucleotide monophosphate (NMP)

binding site (Figure 3.1) (Cui et al., 2013; Lavie et al., 1998; Ostermann et al., 2000b; Li de la Sierra et al., 2001; Caillat et al., 2008; Whittingham et al., 2010; Sinha and Rule, 2017; Biswas et al., 2017). The P-loop (residues 13-20, human numbering), A-loop (residues 179-184), and LID (residues 134-153) are all involved in binding and orientation of the bound ATP and move concomitantly toward the NMP-binding site during the catalytic cycle (Ostermann et al., 2000b).

Most nucleotide monophosphate kinases are assumed to work by a mechanism that is similar to the extensively characterized monomeric AMP kinase (AMPK), though they lack some of the same structural features. Unlike AMPKs, TMPKs contain a much shorter, unstructured LID region similar to that of uridylate kinase. The shorter LID sequence in eukaryotic type I TMPKs does not contain the arginine that directly interacts with the phosphates of ATP as found in type II TMPKs (Lavie et al., 1998; Scheffzek et al., 1996). This suggests that the coordinated motions of the P-loop, A-loop and LID which are unique to type I TMPKs compensate for the reduction in size and lack of a phosphate interacting partner in the shortened LID. Crystallographic structures taken along the reaction coordinate suggest movement of the P-loop, A-loop and LID from the so-called "P-loop open" or "P-loop partially-closed" conformation (obtained with an ATP analog and TMP) to the active "P-loop closed" conformation (obtained with a transition state analog Ap_5dT) is required for catalysis (Ostermann et al., 2000b). Significant structural changes during the catalytic cycle include a 2 Å shift of the mainchain atoms of the P-loop towards the bound TMP and the side chain of Asp15 which swings towards the 3'OH of the bound TMP from an initial distance of 6.8 Å (3'OH to Oδ1) in the TMP ADP complex to 3.7 Å in the Ap_5dT complex. These changes are driven by hydrogen bonds between Asp15 on the P-loop and the side chains of Gln157 on α7 and Arg97 in the NMP-binding site. These interactions are

important for catalytic proficiency since mutation of Asp15 to Ala, Asn, or Glu results in an inactive enzyme (Ostermann et al., 2000b; Brundiers et al., 1999).

3.1 hTMPK, HIV and Cancer

There is considerable interest in the structure and dynamics of human thymidylate kinase (hTMPK) from a therapeutic perspective. Human thymidylate kinase has shown promise as a potential anti-cancer target as its inhibition in LKB1-mutant lung cancer cells is lethal (Liu et al., 2013). Combined inhibition of hTMPK and treatment with doxorubicin impairs DNA repair and reduced tumor cell growth both *in vitro* and *in vivo* (Hu et al., 2012; Chen et al., 2016). In addition to its potential role in cancer treatment, hTMPK catalyzes the rate-limiting step for the activation of antiretroviral thymidine analog prodrug 3'-azido-3'-deoxythymidine (AZT or zidovudine). The slow conversion of AZTMP results in a buildup of the toxic monophosphate form of AZT (Furman et al., 1986; Yan et al., 1995; Lavie et al., 1997; Kline et al., 2009). Structures of the WT enzyme in complex with ADP and TMP or AZTMP show very few differences. Differences between TMP and AZTMP are seen for the catalytically important partially closed state that occurs when TMP and AppNHp are bound in the active site. In the case of TMP the P-loop is drawn closer to the NMP-binding site due to the formation of a hydrogen bond between Asp15H_N and Gln157Oϵ1. This hydrogen bond does not occur in the WT complex with AZTMP and AppNHp (Ostermann et al., 2000b,a). Thus, the low activity towards AZTMP has been attributed to the presence of the bulky azido group that blocks the approach of the P-loop to TMP during catalysis. This hypothesis was supported by structures of hTMPK in complex with a bisubstrate inhibitor where the distance between the Asp15Cγ and C3' on Ap_5dT is 4.2 Å while this distance is 6.7 Å in the Ap_5AZT complex. Structures

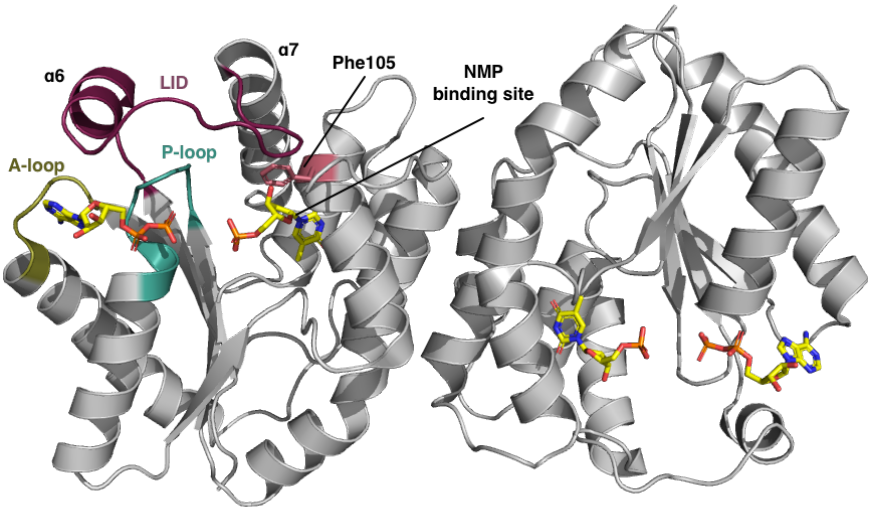


Figure 3.1: Important structural features of hTMPK. Locations of the P-loop (cyan), A-loop (yellow) and LID (red) are indicated on one monomer. These elements move concomitantly toward $\alpha 7$ of the NMP-binding site (labeled) during the catalytic cycle. ADP and TMP shown in yellow, rendered as sticks (PDB 1e2d).

of hTMPK with various TMP-analog prodrugs have shown that the presence of a hydrogen bond donor at the 3' position of TMP also helps to stabilize the closed conformation (Ostermann et al., 2003).

A number of eukaryotic TMPKs are catalytically proficient against AZTMP. In particular, the TMPK from Saccharomyces cerevisiae (ScTMPK) is highly active against AZTMP. Brundiers et al. (1999) changed residues in hTMPK to the match those in ScTMPK in an attempt to increase the catalytic potential of hTMPK to AZTMP. Remarkably, a point mutant of hTMPK, that changes Phe105 to Tyr (F105Y), confers a 20-fold increase in activity toward AZTMP while only reducing activity towards the natural substrate (TMP) 4-fold (Table 3.1). The preferential activity of F105Y toward AZTMP has been exploited as a suicide enzyme for gene therapy cancer treatments (Sato et al., 2013).

Crystallographic structures of F105Y hTMPK in complex with either TMP or AZTMP have shown that the mutation shifts the P-loop to a position that is similar to that found in the complex between Ap_5dT and the WT enzyme. This shift is due, in part, to a new hydrogen bond between Tyr105OH and Asp15H_N in the P-loop. It has been suggested that the shift in the P-loop is the structural basis for the higher activity towards AZTMP (Ostermann et al., 2003). However, the approach of the sidechain of Asp15 to the TMP 3'OH would still be blocked by the azido group when AZTMP is used as the substrate. It is inferred from these structures that the additional hydrogen bond between Tyr105OH and Asp15H_N compensates for the steric clash of the sidechain of Asp15 and 3'N_3 of AZTMP. Unfortunately, the structure of the F105Y protein complexed with either AZTMP and AppNHp or Ap_5AZT have not been reported, thus it is not possible to determine what potential changes might occur in the structure as the AZTMP ATP complex approaches the transition state.

Protein-Ligand	K_M^{NMP} (μM)	K_M^{ATP} (μM)	k_{cat} (s^{-1})
WT-TMP	5^b	5^b	0.73^a
WT-AZTMP	12^b	69^b	0.012^a
F105Y-TMP	4.2^c	$-^d$	0.17^a
F105Y-AZTMP	3.8^c	$-^d$	0.25^a

Table 3.1: Kinetic constants for WT and F105Y hTMPK with TMP and AZTMP. [a]from Lavie et al. 1997, [b]from Ostermann et al. 2003 [c] from Wohrl 2005, [d]Not reported

3.2 CPMG Relaxation Disperson

Both in this study and in the Rho130-polymer conjugate study CPMG relaxation dispersion plays a major role in the characterization of differential dynamics between the various complexes. The CPMG experiment is capable of measuring molecular motions on μs-ms timescale. Conveniently, many important biological processes occur on this timescale including protein folding, large domain movements, and catalysis (Palmer and Koss, 2019). A brief description of chemical exchange phenomena and its application to biomolecular NMR follows.

A typical exchange process for a specifically methyl labeled leucine is schematized in Figure 3.2. Labeled in the figure are three quantities which describe the exchange process: the exchange rate k_{ex}, the chemical shift difference $\Delta\omega$ and equilibrium populations of state A and B p_a/p_b (Kleckner and Foster, 2011). These parameters provide kinetic (k_{ex}), thermodynamic (p_a) and structural ($\Delta\omega$) information about the system and can be used to characterize sparsely-populated, catalytically relevant excited states in enzymes (Hansen et al., 2008a). In this example, the leucine jumps or exchanges between two states, A and B with a rate of $k_{ex} = k_a + k_b$. The equilibrium populations of each state are given as 80% for A and 20% for B. In reality these populations tend to be highly skewed, with p_b usually >95%. Because of this, the chemical shift of the B state, ω_b, is usually not visible in the NMR spectrum. Additionally, the individual chemical shifts for each state may not visible because the rate of exchange is greater than the chemical shift difference ($k_{ex} \gg \Delta\omega$), resulting in a single resonance line at a frequency that is the population weighted average of the individual frequencies. This case is known as the fast exchange regime. The NMR spectrum presented in Figure 3.2 is a case of slow exchange, where $k_{ex} \ll \Delta\omega$ resulting in two observable resonance lines with intensities

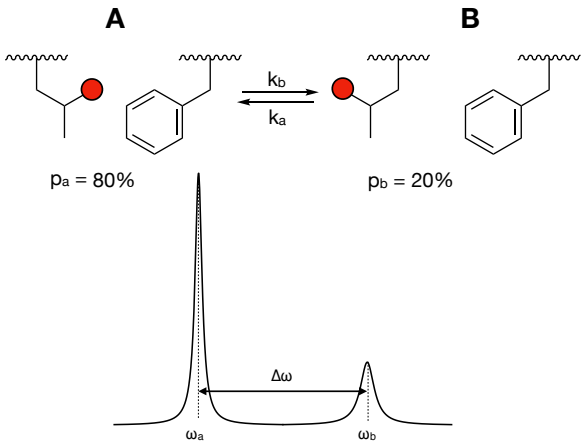


Figure 3.2: Schematic of an exchange process. A leucine residue with the red circles representing a specifically labeled methyl group. The the leucine jumps between two states, A and B with the forward and backward rate constants k_b and k_a, the sum of which is k_{ex} or the exchange rate. These two states produce different chemical shifts with state ω_a being the chemical shift of the A state and ω_b the chemical shift of the B state with difference between them, $\Delta\omega$. The spectrum shown in this figure is for illustrative purposes only and the line widths do not reflect differences in R_2 between the two states. The populations in the A state and B state, respectively, are given as 80% and 20%. The parameters k_{ex}, $\Delta\omega$ and p_a/p_b (exchange parameters) are used to describe a chemical exchange process.

proportional to the populations. Intermediate exchange is the term used to characterize the relationship of k_{ex} to $\Delta\omega$ between these two extremes. In the intermediate exchange regime resonance lines are usually very broad, termed exchange broadened, and may be indistinguishable from noise (Kleckner and Foster, 2011).

Exchange parameters can be extracted from the dispersion curves produced by a CPMG experiment. Chemical exchange is one component of the transverse relaxation rate, R_2:

$$R_2 = R_2^0 + R_{ex} \tag{3.1}$$

where R_2^0 is the transverse relaxation rate in the absence of exchange. The CPMG experiment consists of repeated applications of spin-echos (τ_{cp}-180°-τ_{cp}) which serve to quench the R_{ex} contribution to R_2. The effective transverse relaxation rate ($R_{2,eff}$) is recorded as a function of an increasing frequency of 180° pulses (ν_{CPMG}) over a fixed relaxation delay (T_{relax}). Dispersion curves, the plot of $R_{2,eff}$ versus ν_{CPMG}, are then fit to a mathematical model in order to extract exchange parameters. One such equation, the Carver-Richards equation (Equation 3.2) is valid in all exchange regimes (Carver and Richards, 1972; Lisi and Loria, 2016).

$$\begin{aligned} R_{2,eff} &= R_2^0 + \frac{k_{ex}}{2} - \nu_{CPMG}\cosh^{-1}[D_+\cosh(\eta_+) - D_- cos(\eta_-)] \\ D_\pm &= \frac{1}{2}(\pm 1 + \frac{\Psi + 2\Delta\omega^2}{\sqrt{\Psi^2+\zeta^2}}) \\ \eta_\pm &= \frac{1}{2\sqrt{2}\nu_{CPMG}}\sqrt{\pm\Psi + \sqrt{\Psi^2+\zeta^2}} \\ \Psi &= (\Delta R_2^0 + p_a k_{ex} + p_b k_{ex})^2 - \Delta\omega^2 + 4p_a p_b k_{ex}^2 \\ \zeta &= 2\Delta\omega(\Delta R_2^0 + p_a k_{ex} + p_b k_{ex}) \end{aligned} \tag{3.2}$$

However, in the fast exchange regime, p_a/p_b and $\Delta\omega$ cannot be deconvoluted, preventing structural and thermodynamic information from being extracted

(Kleckner and Foster, 2011). In the limit of fast exchange Equation 3.3, which is considerably simpler, is appropriate (Luz and Meiboom, 1963).

$$R_{2,eff} = R_2^0 + \frac{\phi_{ex}}{k_{ex}}[1 - \frac{\nu_{CPMG}}{k_{ex}} \tanh(\frac{k_{ex}}{4\nu_{CPMG}})] \tag{3.3}$$

Here the term $\phi_{ex} = p_a p_b \Delta\omega^2$. If sufficient information is present, such as dispersion curves collected at multiple fields or curves for multiple residues fit jointly to k_{ex} and p_a, $\Delta\omega$ and p_a can be extracted for processes in fast exchange (Kovrigin et al., 2006). All residues in both Rho130 conjugates and hTMPK are in fast exchange. The one exception is Ile184δ1 in WT hTMPK bound to AZTMP and 2 mM ADP (Section 4.2).

The effect that the different exchange parameters have on the dispersion curves is shown in Figure 3.3. In the simplest case, if there is no chemical exchange process, the curve will be a straight line at R_2. In the fast exchange regime ($k_{ex} \gg \Delta\omega$) a curve resembling simple exponential decay is shown (Figure 3.3 $k_{ex} = 2000$ s^{-1}, black curve top right, and $\Delta\omega = 0.125$ ppm, red curve bottom left). In these cases the simplifications of Equation 3.3 are appropriate. For the other curves in the slow-intermediate exchange regime, other features are present. The relationship of $\Delta\omega$ to k_{ex} determines the prominence of these features. Interestingly, in very slow exchange (Figure 3.3 $k_{ex} = 500$ s^{-1}, red curve top right) the curve has oscillations at lower frequencies. The effect that p_a has on the dispersion curve is to raise and lower its starting point on the R_{ex} axis. Even a 5% difference in p_a (Figure 3.3 $p_a = 95$ %, black curve and 90 %, blue curve) results in a dramatic increase in R_{ex}.

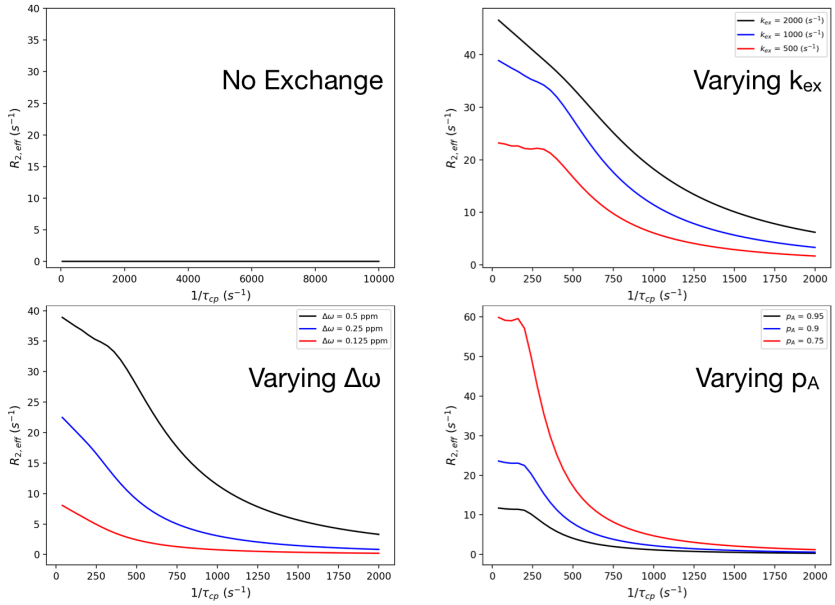


Figure 3.3: Effect of exchange parameters on dispersion curves. (Top Left) An example of No Exchange, a straight line is observed at R_2. (Top Right) Three dispersion curves with different k_{ex}, the three exchange regimes are shown by fixing $\Delta\omega$ and varying k_{ex}: fast (black), intermediate (blue) and slow (red). (Bottom Left) Similar to the varied k_{ex} curves, by varying $\Delta\omega$ the shape of the curve is a simple exponential when $k_{ex} \gg \Delta\omega$. (Bottom Right) Decreasing the population of the ground state, p_a, causes the R_{ex} to increase.

Chapter 4

Dynamics of WT and F105Y hTMPK at Multiple Time Scales

X-ray crystallography studies have provided considerable detail regarding potential conformational changes during catalysis as well as information on the interaction of both the WT enzyme and the F105Y mutant with AZTMP. However, it is still not clear why the F105Y mutant is so proficient at phosphorylating AZTMP. Clearly the dynamics of these important loops play a critical role in transitioning to the catalytically active "P-loop closed" state. Here, we employ spin-relaxation and methyl CPMG relaxation dispersion experiments to better understand the dynamical implications of the 3'-azido group of AZTMP and putative stabilizing effects of the F105Y mutant over multiple time scales. Presented in Section 4.1, ^{13}C HMQC spectra of WT and F105Y hTMPK bound to ADP and either TMP or AZTMP suggest very little difference in the solution structure of the two enzymes. Also, the ps-ns dynamics of these complexes are largely similar, suggesting slow dynamics are likely responsible for the differences in catalytic efficiency. Similar relaxation dispersion studies on AMPK confirmed previous biochemical work and provide insight into the timescales

of transitions into the catalytically active state (Wolf-Watz et al., 2004; Ådén et al., 2012, 2013; Kerns et al., 2015). In Section 4.3 it is shown that AZTMP affects ADP binding in the WT enzyme, leading to destabilization of the closed conformation for the LID and P-loop. Consistent with previous crystal structures, Section 4.4 presents reduced TMP phosphorylation by F105Y hTMPK is due to slow ADP release and a reduced population of the excited state for ATP-binding residues when bound to Ap_5dT. In Section 4.5, we show that the μs-ms dynamics near critical active site residues are highly perturbed in WT hTMPK when bound to AZTMP, while in the F105Y mutant-AZTMP complex this region is stabilized and has similar properties as the WT-TMP complex, explaining the ability of F105Y hTMPK to effectively phosphorylate AZTMP.

4.1 Structural and Dynamical Similarities between WT and F105Y hTMPK

The ^{13}C HMQC spectra of WT and F105Y hTMPK bound to either TMP and ADP (Figure 4.1) or AZTMP and ADP (Figure 4.2) show few CSPs, as anticipated from the similarities in the crystal structures of the four complexes. The assignments of the ILV methyl peaks in the ^{13}C HMQC was performed by Kaustubh Sinha. Although Val14 in the P-loop adopts different positions in the crystal structures it has the same shift in all complexes, suggesting the same average solution structure. Leu57δ1/δ2, Leu161δ1 and Leu167δ1 show the largest, but still modest, CSPs. These residues are proximal to the NMP-binding site and likely reflect local minor structural changes required to accommodate the Phe $\rightarrow$ Tyr mutation (Figure 4.1, inset). Moreover, ^{13}C HMQC spectra of WT and F105Y hTMPK bound to the bisubstrate inhibitor Ap_5dT are essentially identical to their TMP ADP counterparts (Figure 4.2). The largest CSP between the Ap_5dT complex and the TMP or AZTMP complexes

is seen for I184δ1 (0.05 ppm ^{1}H). The lack of significant CSPs in the spectra of the different protein-ligand combinations suggests they all adopt the same average structure in solution.

These spectra clearly indicate that there are changes in the dynamics of the enzyme due to replacement of Phe105 with Tyr. Notably, Figure 4.1 shows that Ile49δ1 is exchange broadened in WT spectra but appears in F105Y spectra. This is quite remarkable since Ile49 is located at the interface between the two subunits and is 16 Å removed from the site of the mutation, suggesting that stabilization of the protein by the F105Y mutation alters the inter-subunit dynamics. Similar effects are seen with the bisubstrate inhibitor Ap_5dT, where the resonances from Ile49 and Ile184 are broader in the WT spectrum (Figure 4.2). The chemical shift of Ile184 is altered by 0.05 ppm in the Ap_5dT spectra, suggesting that the position of the adenine ring in the TMP ADP complex is slightly different than in the Ap_5dT complex; equivalent shifts would be predicted from the crystallographic structures.

The ps-ns dynamics of WT and F105Y hTMPK bound to ADP and either TMP or AZTMP were measured using ^{13}C R_1 and $R_{1\rho}$ relaxation experiments. The S^2 values obtained from analysis of the R_1 and R_2 rates are largely similar for each of the four complexes (Figures 7 and S2). The nature of the bound ligand, TMP or AZTMP, does not affect the order parameters (Figure 4.4). The color mapping shown in Figure 4.3 illustrates the general summary that residues in the β-sheet below the active site are relatively rigid as is the P-loop near the terminal phosphate on ADP. The helix that is amino-terminal to the LID is mobile on the ps-ns time scale scale (S^2 = 0.23 for Leu137). Unfortunately, there are no ILV residues on the unstructured portion of the LID, preventing characterization of its motion. Although the order parameters are similar between WT and F105Y, Val80γ2 has a higher S^2 of 0.76 ± 0.01 in F105Y-AZTMP complex as compared to 0.70 ± 0.03 in the WT-AZTMP

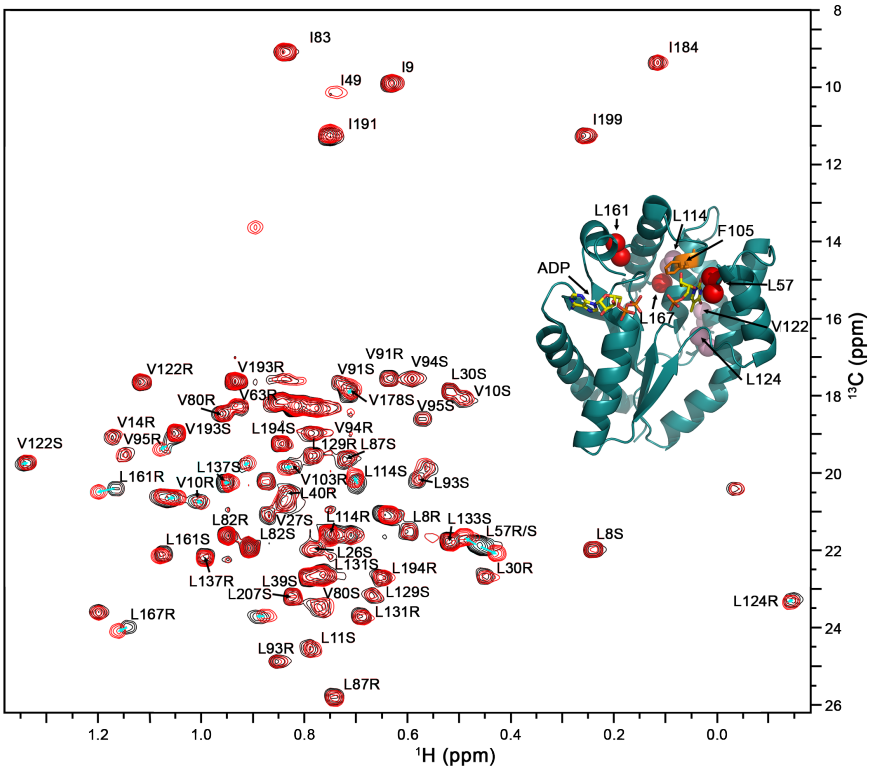


Figure 4.1: Overlay of WT and F105Y hTMPK bound to TMP and ADP ^{13}C HMQC spectra. The WT (black) and F105Y (red) spectra have few CSPs (cyan arrows). The insert shows the location of residues with large (red spheres) and small CSPs (pink spheres) in the structure. The largest CSP was associated with L161R (^{1}H CSP = 0.035ppm, ^{13}C CSP = 0.08ppm) Stereochemistry indicated by R and S; R indicates $\gamma 1$ (V) or $\delta 1$ (L), S indicates $\gamma 2$ (V) or $\delta 2$ (L).

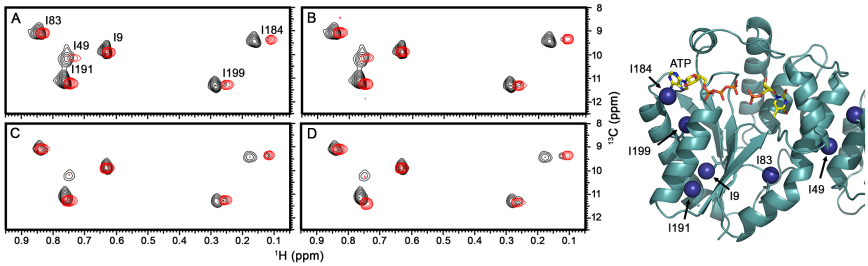


Figure 4.2: ^{13}C HMQC show minor Ap_5dT degradation. F105Y Ap_5dT spectra (top black) overlaid with (A) F105Y TMP ADP and (B) F105Y AZTMP ADP (red). WT Ap_5dT spectra (bottom black) overlaid with (C) WT TMP ADP and (D) WT AZTMP ADP (red). Degradation peaks appear in the same locations as ADP-bound complexes. Binding of Ap_5dT perturbs the chemical shift of I184 to the largest extent, followed by I199, and I191.

complex (Figure 4.3). The enhanced ordering of Val80 is also seen in the TMP complex (Figure 4.4C). This residue is at the dimer interface and its restricted dynamics in the F105Y enzyme parallels our observations from Ile49 which also becomes less dynamic in the F105Y enzyme. Changes in the dynamics of the protein due to replacement of Phe105 with Tyr are propagated to the dimer interface.

Among TMPKs, hTMPK is a particularly slow enzyme with turnover rates nearly an order of magnitude below other eukaryotic counterparts (Sinha and Rule, 2017). Microbial TMPKs are usually adept at AZTMP phosphorylation and share a tyrosine at position 105 (human numbering) which presumably performs the same function of stabilizing P-loop closure in F105Y hTMPK. ScTMPK which has been co-crystallized with TMP alone shows the highly conserved Asp15 on the P-loop contacts Arg97 and the 3'OH of TMP forming the so-called "P-loop closed" conformation and suggesting that even in the absence of ATP, the enzyme exchanges between the open and closed states. This could not be assessed in hTMPK due to stability issues when no ligand is bound. Co-crystallization of the yeast enzyme with Ap_5dT induces similar shifts to those seen in hTMPK bound to Ap_5dT with movements in the A-loop and ordering of the LID, which is disordered in the TMP-bound structure (Lavie et al., 1998). Thus, the biasing toward P-loop closure by F105Y hTMPK may be an artifact observable only at cryogenic temperatures. Indeed, our higher temperature NMR spectra contain few CSPs in the P-loop, A-loop and LID, suggesting no difference in the average solution structure. Similarly, the dynamics of these complexes show no major differences on the ps-ns time scale suggesting slower processes such as larger domain motions may be responsible for catalytic differences.

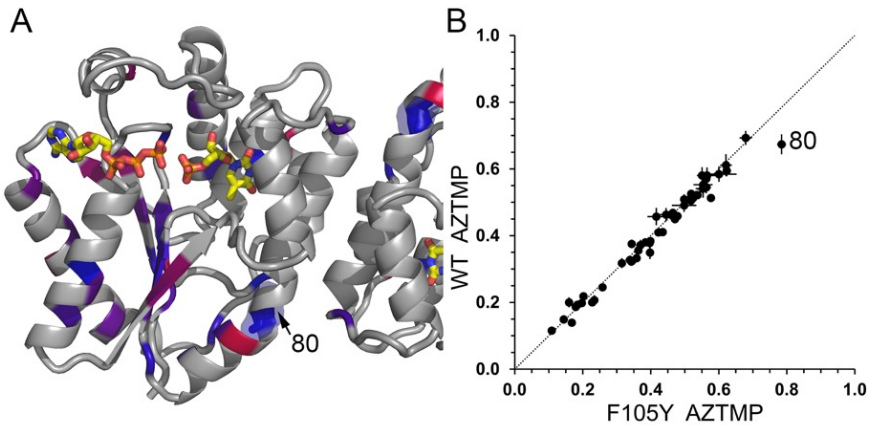


Figure 4.3: F105Y induced changes to ps-ns dynamics at the dimer interface. (A) Mapping of S^2 order parameters onto the X-ray structure using a blue-to-red color map. Blue represents high order parameters and red presents low order parameters. Val80γ2 (labeled, rendered in sticks and highlighted) shows an increased S^2 for the F105Y mutant. (B) Correlation plot of order parameters for WT and F105Y bound to AZTMP and ADP show significant differences in S^2 for only Val80γ2.

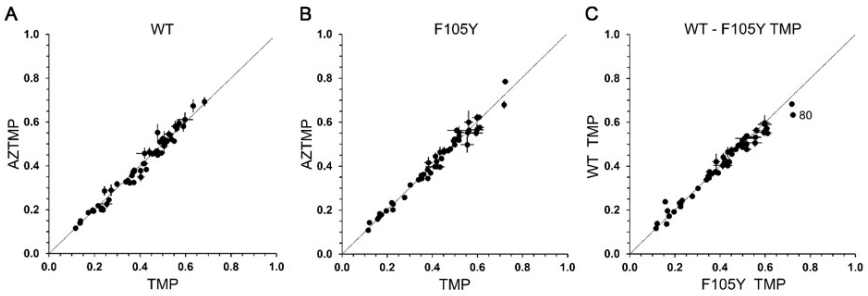


Figure 4.4: Methyl order parameters for WT and F105Y in the presence of TMP and AZTMP. (A) and (B) High frequency (ps-ns) motions are not affected by the nature of the bound NMP (TMP versus AZTMP). (C) Most of the methyl groups have the same order parameter in the TMP complex of WT and F105Y, except for Val80, which is more ordered in the F105Y complex.

4.2 Ile184 is a sensitive probe of ADP saturation and dynamics

Ile184δ1 undergoes a dramatic ^{1}H CSP of 0.72 ppm upon addition of ADP to hTMPK bound to either TMP or AZTMP (Figure 4.5A). This can be attributed to the ring current shift (Perkins and Wüthrich, 1979) experienced by Ile184, which is in close proximity with the aromatic adenine ring of ADP (Figure 4.5B). Relaxation dispersion measurements for WT hTMPK with bound AZTMP at two concentrations of ADP illustrate the high sensitivity of relaxation dispersion to identify sparsely populated states. Because of the large $\Delta\omega$ between the ADP-bound and ADP-free states, R_{ex} ($=p_a p_b \Delta\omega^2/k_{ex}$) shows a marked increase as the population of the free state, p_b, at the lower ADP concentration. The two Ile184δ1 dispersion curves for ADP concentrations of 2 mM and 8 mM result in fitted exchange parameters describing two different processes (Figure 4.5C).

Because Ile184δ1 is in slow exchange in the WT-AZTMP 2mM ADP complex its dispersion curve could be reliably fit to separate p_a and $\Delta\omega$, allowing for a comparison to the observed CSP for ADP binding and the binding parameters obtained from the ADP titration. The fitted Ile184δ1 $\Delta\omega$ of 0.62 $\pm$ 0.014 ppm with 2 mM ADP bound (Table 4.1) is consistent with the observed CSP of 0.72 ppm, while the $\Delta\omega$ of 0.35 $\pm$ 0.05 ppm with 8 mM ADP when fit to the full model (not shown) is not consistent. Likewise, the p_a of 98.6 $\pm$ 0.08% for the WT-AZTMP 2mM ADP complex is in close agreement with the theoretical fractional saturation of 98.8%. Fractional saturations of protein-ligand complexes were calculated by:

$$\theta = \frac{[PL]}{[P_{tot}]} = \frac{(K_D + [P_{tot}] + [L_{tot}]) - \sqrt{(K_D + [P_{tot}] + [L_{tot}])^2 - 4[P_{tot}][L_{tot}]}}{2[P_{tot}]} \tag{4.1}$$

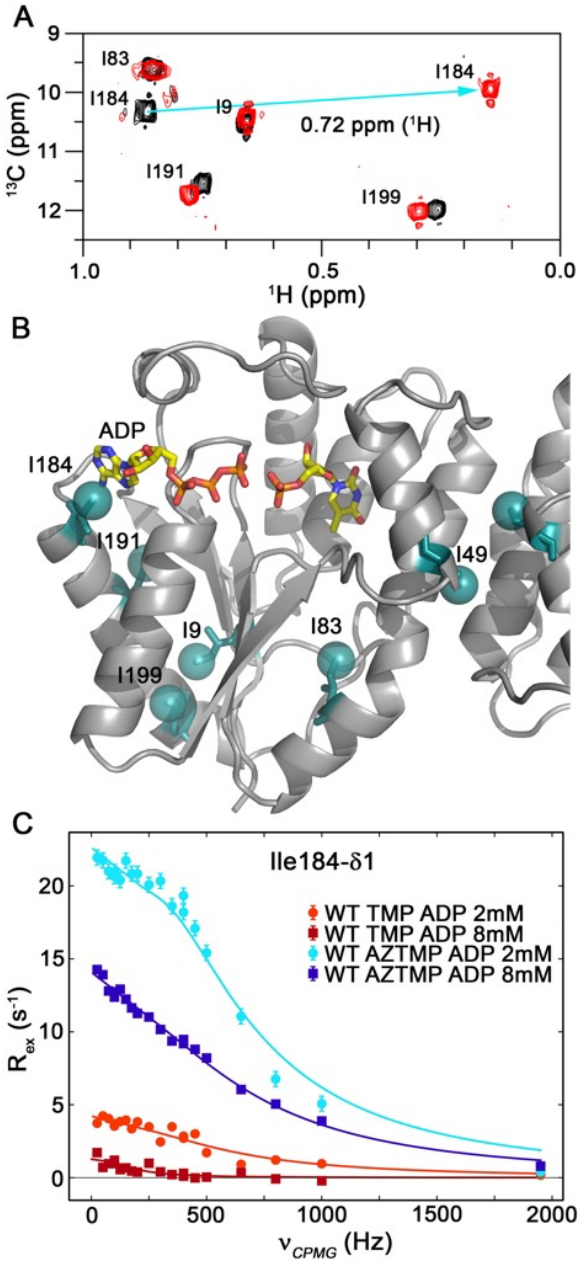


Figure 4.5: Ile184δ1 is sensitive to adenosine nucleotide saturation. (A) Overlay of WT hTMPK bound to TMP (black) and TMP and ADP (red) results in a ^{1}H CSP of 0.72 ppm. (B) The interaction of Ile184 with the adenine ring of ADP causes a ring current shift. (C) Dispersion curves (850 MHz) for WT hTMPK bound to TMP (red curves) or AZTMP (blue curves) and two different concentrations of ADP (2mM lighter; 8mM darker). R_2^0 is subtracted from the data, normalizing the curve to 0 to facilitate comparisons.

as the concentrations of ligands were often not in large excess so [PL] made up a large portion of [L_{tot}].

Estimates of K_D and k_{off} from CPMG relaxation dispersion were calculated from p_a and k_{ex}. The bound ligand concentration is calculated assuming p_a = [PL]/[P_{tot}], and each sample has a known amount of protein. The free ligand concentration is estimated as [L] = [L_{tot}] – [PL]. An estimate of K_D can be made from the individual rate constants given the exchange process:

$$A \underset{k_a}{\overset{k_b}{\rightleftharpoons}} B$$

and $k_{ex} = k_a + k_b$ where $k_a = p_b k_{ex}$ and $k_b = p_a k_{ex}$. Assuming the process is a ligand binding event,

$$P + L \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} PL$$

then by analogy $k_{ex} = [L]k_{on} + k_{off}$, where the first order rate constant k_{off} = k_a and the second order rate constant $k_{on} = k_b/[L]$, where [L] is the free ligand concentration. Using the calculation for [L] above, $K_D = k_{off}/k_{on} = k_a[L]/k_b$. This method was used on parameters fit to Ile184δ1 for the WT AZTMP protein with 2 mM ADP. The estimated binding parameters were: $K_D = 17 \pm 4.0$ μM and $k_{off} = 31 \pm 5.1$ s^{-1} for ADP binding to WT AZTMP. This is consistent with the K_D of 15 ± 1 μM and k_{off} of 24 ± 0.8 s^{-1} obtained from the ADP titration experiment.

In order to capture relevant motions contributing to ligand-induced conformational change of the P-loop, LID and A-loop on the μs-ms timescale it is critical to ensure that contributions to R_{ex} from ligand exchange with the free pool is effectively zero. For the following, we limit our discussion to complexes with 8 mM ADP to ensure saturating conditions of ADP and dispersion measurements which report on internal dynamics of the protein and not simply ADP exchange. Relaxation dispersion of Ile184 shows that the adenine ring is effectively static on the μs-ms timescale in the F105Y complex with either

Protein-Ligand	k_{ex} (s^{-1})	p_a (%)	$\Delta\omega$ (ppm)	ϕ_{ex} ($rad^{-2}.s^{-2}$)	R_2^0 (s^{-1})	χ^2/N
WT-TMP-2 mM ADP	4000 ± 440	-	-	18000 ± 6100	12 ± 0.39	2.9
WT-TMP-8 mM ADP	1300 ± 240	-	-	2400 ± 560	13 ± 0.12	2.2
WT-AZTMP-2mM ADP[a]	2200 ± 360	98.6 ± 0.08	0.620 ± 0.014	-	14.7 ± 1.5	4.2
WT-AZTMP-8 mM ADP	3800 ± 100	-	-	56000 ± 2000	13 ± 0.13	6.9

Table 4.1: Effect of ADP concentration on exchange parameters for WT Ile184δ1. [a]Fit to the full model, Equation 3.2 Fitted parameters for F105Y are omitted here because there is essentially no exchange in either the TMP or AZTMP complex, regardless of the concentration of ADP, due to the smaller K_D for the binding of ADP to the F105Y protein, leading to full saturation.

TMP or AZTMP bound. At an ADP concentration of 2 mM (not shown) or 8 mM (Figure 4.6A), both F105Y-TMP and F105Y-AZTMP showed a $R_{ex} \approx 0$ and could not be reliably fit to an exchange model, suggesting that there no exchange process for Ile184δ1 in F105Y. In contrast, both WT complexes undergo a fast exchange process, even when saturated with ADP. This process is approximately 3-fold faster in WT-AZTMP with a 20-fold increase in ϕ_{ex} (Table 4.1), indicating significant motion of the adenine ring in the WT-AZTMP complex. This change is likely due to changes in both p_b and $\Delta\omega$ to account for the large increase in ϕ_{ex}.

4.3 AZTMP affects adenosine nucleotide handling in WT hTMPK

Val14, Leu137 and Leu161 are in close contact in the closed form of the enzyme, bringing together the P-loop, LID and α7 in the NMP binding site. These three methyl groups serve as ideal probes to monitor changes to those regions involved in the shift from the "P-loop open" to the active "P-loop closed" conformations. The dispersion curves show: WT-AZTMP has increased R_{ex} compared to WT-TMP, F105Y-AZTMP has a slightly reduced R_{ex} and F105Y-TMP has a similar or lower R_{ex} than F105Y-AZTMP. The fitted parameters, summarized in Tables 4-7, show a general trend in that the exchange rates for WT-TMP and WT-AZTMP are equivalent or higher than F105Y, and ϕ_{ex} is larger for the WT protein.

The dynamics of the P-loop, as reported by Val14γ1, are related to the catalytic efficiency of each complex. The k_{ex} in the WT-TMP complex is largest with the F105Y complexes being almost halved (Figure 4.6B). Notably both F105Y complexes were fit to nearly the same exchange rate. The least efficient complex, WT-AZTMP shows a lower exchange rate and a higher value

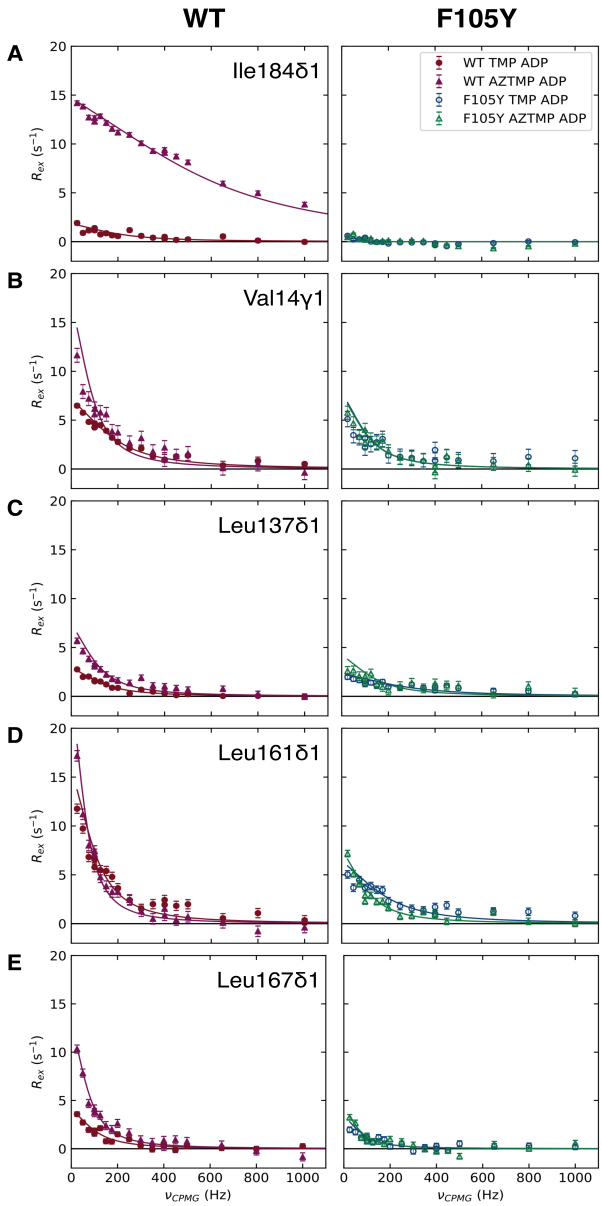


Figure 4.6: 850 MHz dispersion curves for residues near the ATP-binding site.(Left) WT complexes with TMP ADP (circles) and AZTMP ADP (triangles). (Right) F105Y complexes with TMP ADP (open circles) and AZTMP ADP (open triangles). Solid lines represent a best fit to the fast exchange model (Equation 3.3) for all residues except F105Y Ile184δ1 which were fit to a "no exchange" model. R_2^0 is subtracted from the data, normalizing the curve to 0 to facilitate comparisons. WT-AZTMP ADP shows increases to R_{ex} across all of these residues, but F105Y-AZTMP does not.

of ϕ_{ex}, suggesting a higher occupancy of an unproductive state. The ϕ_{ex} values for WT-TMP and WT-AZTMP are approximately double those seen for the F105Y, indicating that the F105Y mutation likely reduces the population of the minor state (Table 4.2).

Leu137 is contained in the α6 of the LID. The 2-fold increase in ϕ_{ex} for Leu137δ1 in the WT-AZTMP complex versus the WT-TMP complex suggest some destabilization of the LID in WT-AZTMP (Table 4.3). Also, the k_{ex} for F105Y-TMP is notably higher than the other complexes pointing toward a more rapid interconversion between the major and minor conformers. The exchange rate of F105Y-AZTMP is similar to that of WT-TMP.

The dynamics of the helix that follows the disordered portion of the LID (α7), as reported by Leu161δ1, show a similar trend as Leu137δ1, in that the exchange rate of the F105Y-TMP complex is higher than the other three complexes (Table 4.4). The reduced k_{ex} for WT-AZTMP may reflect the unfavorable transition to the closed conformation. Also, the ϕ_{ex} is distinct between the mutant and WT protein with a slight increase for the F105Y-TMP complex.

Leu167δ1, positioned at the C-terminal end of α7, is fit to a similar k_{ex} for all four complexes. WT-AZTMP has a 2-fold increase in ϕ_{ex} compared to WT-TMP. The ϕ_{ex} for both F105Y-TMP and F105Y-AZTMP is more similar to WT-TMP, meaning this dynamical perturbation is unique to WT-AZTMP (Table 4.5).

Differences on the μs-ms time scale due to transition between the open and closed conformation, were revealed by methyl relaxation dispersion experiments as dynamical perturbations to the residues on the P-loop (Val14γ1), LID (Leu137δ1) and α7 near the active site (Leu161δ1). Our overall framework for analysis is that the excited state is the closed form of the enzyme. In the case of the WT enzyme complexed with AZTMP this structure is not conducive to phosphate transfer while in the case of the F105Y enzyme a different tra-

Protein-Ligand	k_{ex} (s^{-1})	ϕ_{ex} ($rad^{-2} \cdot s^{-2}$)	R_2^0 (s^{-1})	χ^2/N
WT-TMP	1200 ± 56	8600 ± 340	14 ± 0.09	4.2
WT-AZTMP	590 ± 110	10000 ± 710	15 ± 0.23	7.2
F105Y-TMP	730 ± 72	5700 ± 360	15 ± 0.14	3.4
F105Y-AZTMP	740 ± 58	5600 ± 280	15 ± 0.09	1.2

Table 4.2: Exchange parameters for Val14γ1

Protein-Ligand	k_{ex} (s^{-1})	ϕ_{ex} ($rad^{-2} \cdot s^{-2}$)	R_2^0 (s^{-1})	χ^2/N
WT-TMP	880 ± 48	2600 ± 70	9.7 ± 0.02	2.3
WT-AZTMP	730 ± 150	5500 ± 620	9.3 ± 0.14	5.5
F105Y-TMP	1700 ± 270	4000 ± 640	9.7 ± 0.10	2.4
F105Y-AZTMP	910 ± 140	3900 ± 440	8.8 ± 0.12	42

Table 4.3: Exchange parameters for Leu137δ1

Protein-Ligand	k_{ex} (s^{-1})	ϕ_{ex} ($rad^{-2} \cdot s^{-2}$)	R_2^0 (s^{-1})	χ^2/N
WT-TMP	700 ± 33	11000 ± 340	12 ± 0.14	8.2
WT-AZTMP	430 ± 54	10000 ± 310	14 ± 0.14	3.3
F105Y-TMP	1200 ± 120	7900 ± 570	12 ± 0.12	4.6
F105Y-AZTMP	730 ± 110	5600 ± 610	12 ± 0.15	5.1

Table 4.4: Exchange parameters for Leu161δ1

Protein-Ligand	k_{ex} (s^{-1})	ϕ_{ex} ($rad^{-2} \cdot s^{-2}$)	R_2^0 (s^{-1})	χ^2/N
WT-TMP	690 ± 92	2900 ± 220	12 ± 0.06	1.9
WT-AZTMP	500 ± 34	6500 ± 190	14 ± 0.10	1.8
F105Y-TMP	540 ± 130	1800 ± 250	12 ± 0.10	3.8
F105Y-AZTMP	460 ± 52	2000 ± 300	12 ± 0.07	2.3

Table 4.5: Exchange parameters for Leu167δ1

jectory is taken such that the enzyme can facilitate phosphate transfer. The dynamics of Val14γ1 partially explains previous crystallographic and kinetic data. Because the ϕ_{ex} values are similar for both WT-TMP and WT-AZTMP (8,600 versus 10,000 rad^2·s^2), it is reasonable to assume the same motion is captured in both complexes. The exchange rate is halved for WT-AZTMP, dropping from 1200 to 590 s^{-1}, indicating that AZTMP blocks the approach of the P-loop toward the NMP-site, the "P-loop closed" excited state. Both F105Y-TMP and F105Y-AZTMP have nearly identical exchange rates of 730 and 740 s^{-1}, both of which are faster than WT-AZTMP. This is taken to show that the F105Y mutation allows the enzyme to access a closed form regardless of substrate in the NMP site. The modest decrease in ϕ_{ex}, as well as k_{ex} in the F105Y complexes reflects the reduced population and exchange to the closed state, which may explain the overall reduced activity of the mutant. Leu137δ1 and Leu161δ1 report on helical movements near the active site and show distinct dynamics which are both ligand and protein dependent. We speculate that these residues, situated at the interface of the ATP- and NMP-binding sites, are sensitive to P-loop closure. The exchange rate for WT-AZTMP is reduced relative to WT-TMP while the rate for F105Y-TMP is somewhat faster. Interestingly, the exchange rates for WT-TMP and F105Y-AZTMP are similar, suggesting a requirement for a particular frequency for efficient catalysis. The possibility of Mg^{2+} exchange contributing to the observed relaxation dispersion was considered, however it seems an unlikely cause for our results. The binding studies, as well as evidence from previous crystal structures suggest differences in ADP handling between the WT and F105Y proteins, which for WT is dependent on which nucleotide occupies the NMP-site.

Because of the large ring current effect, Ile184δ1 serves as a powerful probe of ADP saturation and adenosine ring dynamics. The presence of a fast exchange process in saturating conditions of ADP suggests that the adenine ring

is mobile within the WT enzyme. The WT-AZTMP complex has a 3-fold increase in k_{ex} as well as a 20-fold increase in ϕ_{ex}, suggesting the adenine ring is quite mobile compared to the WT-TMP complex. In addition to Ile184, the adenine ring is held in place by a cation-π interaction with Arg143 in the LID on the C-terminal end of α6 which contains Leu137δ1 (Figure 4.7). This leucine shows an increased ϕ_{ex} in WT-AZTMP, consistent with a dynamical perturbation in this region that may be mediated by Arg143. It is clear from both the relaxation dispersion parameters and 2-fold increase in k_{off} for ADP with AZTMP bound, that ADP handling is different in the WT-AZTMP complex than in the TMP complex. This may involve a disruption in communication between these regions that is mediated by the guanidinium group of Arg143, which interacts with the adenine ring as well as 2 hydrogen bonds to the backbone carbonyls of Arg16 (P-loop) and Ala180 (A-loop) (Figure 4.10). Likewise, perturbations to P-loop dynamics may propagate to the LID through the sidechain of Arg16.

4.4 Catalytic rate of F105Y hTMPK is reduced because of altered ADP Binding and transition state dynamics

We directly measured the binding constants for ADP in the presence of either TMP or AZTMP. This was achieved by two-dimensional line shape fitting of resonance lines with CSPs from a series of ^{13}C HMQC spectra with increasing concentrations of ADP titrated into WT or F105Y hTMPK bound to either TMP or AZTMP (Table 4.6). The K_D is not dependent on the identity of the ligand in the NMP site, however F105Y hTMPK has a 2.5-fold higher affinity for ADP compared to WT. Interestingly, the k_{off} is ligand dependent for both proteins, with AZTMP showing a 2-fold faster off-rate with the WT

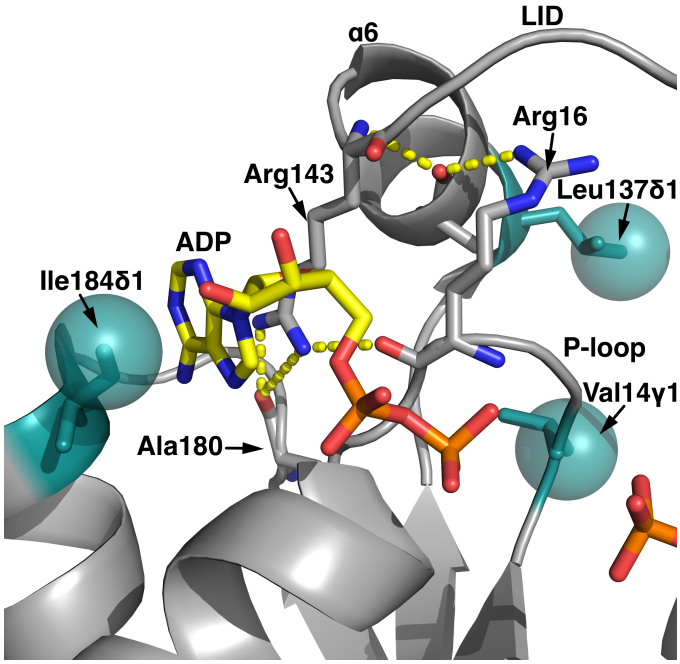


Figure 4.7: Dynamic events in the adenosine binding site are propagated to the LID by two arginines. ILV methyl probes (cyan sticks and spheres, labeled) report on the coordinated motions in the ATP-binding site. The ADP binding site contains three important loop regions: the P-loop (reported on by Val14γ1), A-loop (reported on by Ile184δ1) and the LID (reported on by Leu137δ1). In WT-AZTMP P-loop closure is prevented causing adenosine to be mishandled resulting in a highly mobile adenine ring. These dynamical perturbations are propagated to the LID through Arg143 and Arg16.

protein and a 15-fold reduced off-rate in the case of F105Y-AZTMP. The faster off-rate of ADP from the WT-AZTMP complex may reflect the unfavorable interaction between the 3'-azido group and Asp15 on the P-loop.23 The off-rate of ADP from F105Y-AZTMP is similar to the catalytic rate (0.25 s-1, Table 3.1), suggesting that product release may be rate-limiting.

While the F105Y mutant is supposedly predisposed to enter the catalytically active "P-loop closed" state, it does have a reduced catalytic rate. The ADP titration suggests that F105Y hTMPK has increased affinity for ADP, and Ile184δ1 shows no exchange for both 2 mM and 8 mM samples of either F105Y complex. Human TMPK has been shown to follow a random bi-bi mechanism suggesting that the order of product release is random, however a recent random accelerated molecular dynamics study of WT hTMPK suggests that TDP release pathways are largely biased toward the residues immediately following α6 in the LID whether or not ADP had already dissociated (Chaudhary et al., 2019). This suggests ADP release may be responsible for some loss in activity as this region is stabilized in the F105Y complexes, reported on by Leu137δ1.

To determine the dynamics of WT and F105Y hTMPK as they approach the transition state, CPMG relaxation dispersion experiments were performed on complexes of these enzymes to the bisubstrate inhibitor Ap_5dT. Crystallographic structures have shown that Ap_5dT induces the fully closed conformation in WT hTMPK.8 Our relaxation dispersion experiments show that WT hTMPK shows enhanced exchange in the presence of Ap_5dT for residues in the LID region (Leu137δ1), A-loop (Ile184δ1) and NMP-binding site (Leu161δ1, Leu167δ1) (Figure 4.9), suggesting intrinsic motions in these regions may be important for formation of the transition state. The increase in exchange for these residues is attenuated in F105Y-Ap_5dT complex (Figure 4.9). A comparison of the exchange for WT-Ap_5dT and F105Y-Ap_5dT (Figure 4.8) shows

Protein-Ligand	K_D (μM)	k_{off} (s^{-1})
WT-TMP	15 ± 2.1	12 ± 0.7
WT-AZTMP	15 ± 1.3	24 ± 0.8
F105Y-TMP	4.2 ± 1.3	2.7 ± 0.8
F105Y-AZTMP	4.4 ± 1.7	0.2 ± 0.9

Table 4.6: ADP binding constants determined by two-dimensional lineshapes

that in all cases the exchange for F105Y is reduced compared to the WT complex and the differences between the two proteins is largest for residues in the P-loop (Val14), LID (Leu137), A-loop (Ile184) and the NMP binding site (Leu161, Leu167). For most residues, with the exception of Leu161, the k_{ex} was similar in both proteins and the WT-Ap_5dT complex showed higher values of ϕ_{ex} (Table S1).

A second mechanism for loss of activity is indicated by the exchange parameters for the F105Y Ap_5dT complex which show a reduced population of the minor state across the residues in the ATP-binding site (Val14γ1, Leu137δ1, Leu161δ1, Ile184δ1) as compared to WT. Since Ap_5dT induces the closed form of the enzyme, we assign the excited state to a previously uncharacterized catalytically required intermediate. Because the F105Y mutation reduces conformational motion in the ATP-binding site it is less favorable to encounter this intermediate, resulting in a reduced catalytic rate. The combined effects of slow product release and a less populated active state explain the loss in activity for F105Y hTMPK.

The appearance of Ile49δ1 in the Ap_5dT ^{13}C HMQC spectra may report on the dynamics of the NMP-site in the catalytically productive state. Ile49 resides at the dimer interface proximal to Arg45 which interacts with the α-phosphate of the TMP possibly reporting on inter-subunit communication between the adjacent NMP-sites or an active conformation for Arg45 (Figure 4.10). Likewise, appearance of this resonance in the F105Y spectra may indicate the stabilization of this same state. It has been shown using a chemical biology approach that the rotation of the α-phosphate away from Arg97 and toward Arg45 is required for efficient catalysis of ribose-modified TMP nucleotides.24 The effect of the transition state complex Ap_5dT on the dynamics of the WT enzyme shows increased exchange for Leu137δ1 and Leu161δ1 when compared to the TMP complex (Figure 4.9). In contrast, these two residues show similar

	WT Ap_5dT				F105Y Ap_5dT			
Residue	k_{ex} (s^{-1})	ϕ_{ex} ($rad^{-2}.s^{-2}$)	R_2^0 (s^{-1})	χ^2/N	k_{ex} (s^{-1})	ϕ_{ex} ($rad^{-2}.s^{-2}$)	R_2^0 (s^{-1})	χ^2/N
Val10γ2	830 ± 87	13000 ± 1200	20 ± 0.34	0.88	800 ± 43	8700 ± 370	18 ± 0.12	1.8
Val14γ1	1000 ± 62	11000 ± 580	17 ± 0.15	3.7	1500 ± 180	6400 ± 810	16 ± 0.15	2.2
Val95γ2	500 ± 27	9300 ± 350	20 ± 0.15	2.0	360 ± 23	6600 ± 270	18 ± 0.15	1.5
Leu137δ1	770 ± 40	11000 ± 420	10 ± 0.14	1.6	580 ± 100	2600 ± 320	11 ± 0.10	0.45
Leu161δ1	1400 ± 87	39000 ± 2500	15 ± 0.50	6.7	580 ± 220	5300 ± 1300	15 ± 0.26	7.4
Leu167δ1	950 ± 45	25000 ± 1000	14 ± 0.24	2.6	800 ± 56	7000 ± 410	12 ± 0.14	1.5
Ile184δ1	2000 ± 180	92000 ± 9000	22 ± 1.37	8.4	2400 ± 260	35000 ± 5000	16 ± 0.71	0.99

Table 4.7: Exchange parameters for hTMPK Ap_5dT complexes. Data was acquired at a single static field (850 MHz). Errors were estimated using a Monte Carlo method.

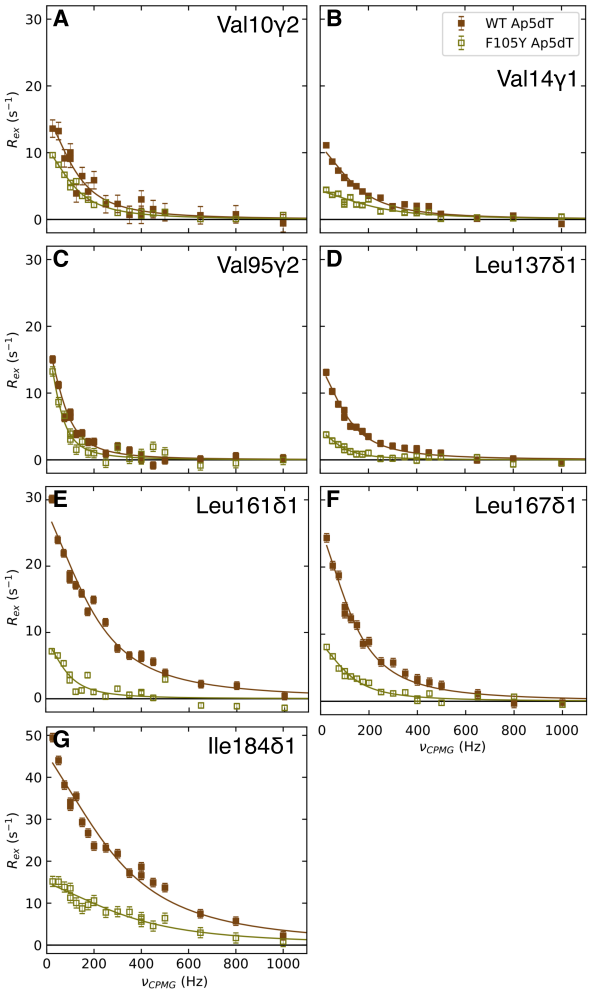


Figure 4.8: Ap_5dT increases dynamics of WT enzyme for residues near the ATP-binding site. Dispersion curves acquired at 850 MHz are shown for the WT complexes with Ap_5dT (brown squares) and F105Y complex with Ap_5dT (yellow open squares). R_2^0 is subtracted from the data, normalizing the curve to 0 to facilitate comparisons. N.B. the maximum of the y-axis is 50 s^{-1} for Ile184. Solid lines represent a best fit to the fast exchange model (Equation 3.3) for all residues. A reduced population of the minor state in the F105Y complex reduces the R_{ex}. Dynamical differences in the approach of F105Y hTMPK into the transition state may contribute to the enzyme's reduced activity with TMP.

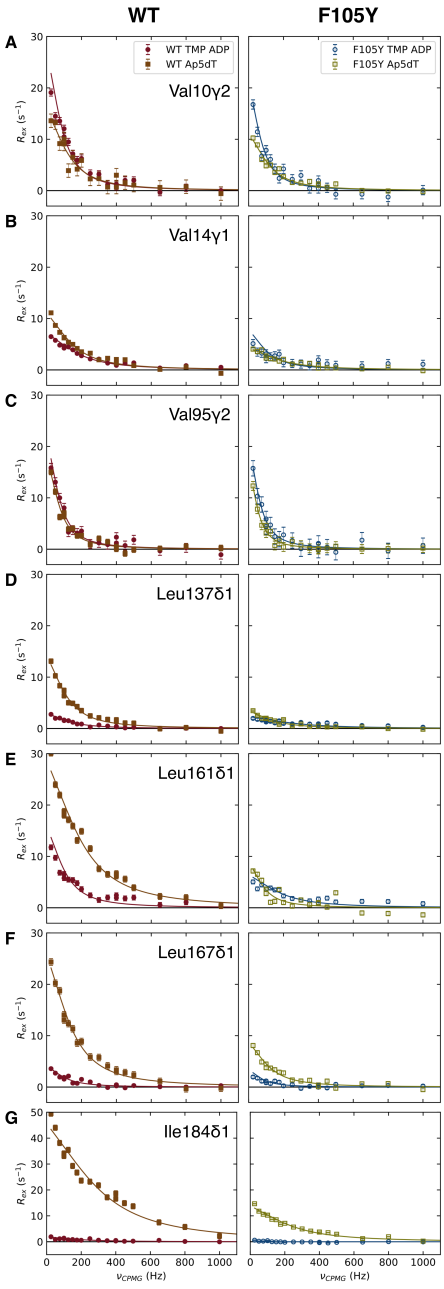


Figure 4.9: Comparison of hTMPK Ap_5dT and TMP dispersion curves. The TMP curves from Figure 4.6 and Ap_5dT curves from 4.8 are overlaid for comparison. The F105Y curves have a decreased R_{ex} compared to WT, suggesting a reduced p_b.

exchange rates for the F105Y mutant, suggesting that the stabilizing effect of Tyr105 extends to Ile49 via Arg45. This partially explains the similar catalytic rates observed for both TMP and AZTMP by F105Y hTMPK.

4.5 WT-AZTMP has altered dynamics near the catalytic center

Exchange properties for Val10γ2 and Val95γ2 reveal another region of the protein affected by the ligand bound in the NMP site and the F105Y mutation, the interior β-sheet. Val10 and Val95 are situated on adjacent β-strands (β1 and β3, respectively) at the interface between the P-loop and catalytically important DRX motif (Figure 4.11B).8,13 In contrast to the residues near the ATP-binding site, this pair shows a reduced R_{ex} for WT-AZTMP, increased R_{ex} for WT-TMP, and an intermediate R_{ex} for both F105Y complexes (Figure 4.11A). The WT-TMP, F105Y-TMP and F105Y-AZTMP complexes all show essentially the same k_{ex} of approximately 550 s^{-1} for both residues. In contrast, the WT-AZTMP complex has an increased k_{ex} for both residues. Interestingly for Val95γ2, WT-TMP and F105Y-AZTMP have nearly identical exchange parameters, while both WT-AZTMP and F105Y-TMP, the substrates with reduced catalytic rate, show larger and smaller values for k_{ex}, respectively (Table 4.8). Both WT-AZTMP and F105Y-TMP have a reduced ϕ_{ex} as well, suggesting the minor state that is more readily accessible in the WT-TMP and F105Y-AZTMP. The exchange rates for Val10γ2 have a similar trend to Val95γ2. (Table 4.9).

The exchange parameters for Val95γ2, a residue adjacent to the critical DRX motif (residues 96-98), are also indicative of dynamical perturbations related to the catalytic rate. The WT-AZTMP complex has a 1.5-fold increased k_{ex} for both Val10γ2 and Val95γ2, while the more catalytically competent

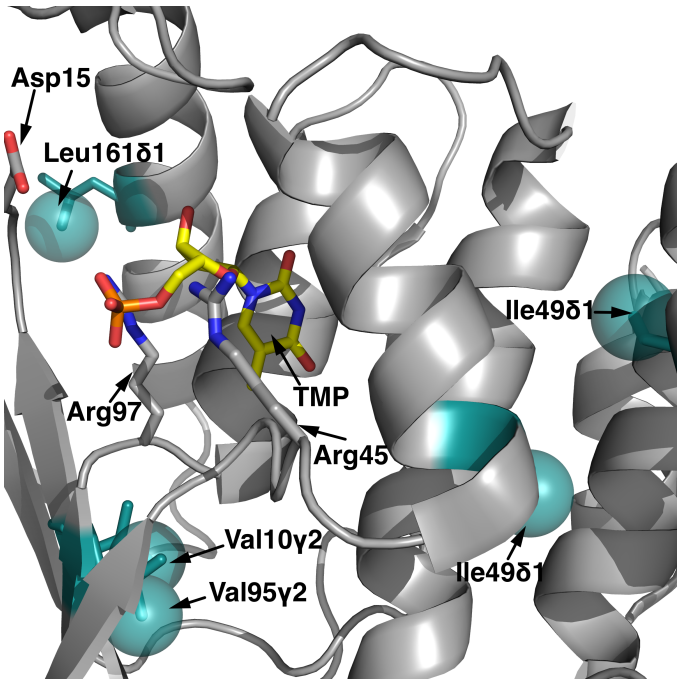


Figure 4.10: Reduction in exchange of Ile49 in F105Y is due to the stabilization of Arg45 in the closed complex. The guanido group of Arg45 interacts with the α-phosphate of dTMP and Glu149 (not shown) in the LID. This interaction also stabilizes Leu161δ1.

Protein-Ligand	k_{ex} (s^{-1})	ϕ_{ex} ($rad^{-2} \cdot s^{-2}$)	R_2^0 (s^{-1})	χ^2/N
WT-TMP	500 ± 29	11000 ± 340	20 ± 0.13	1.6
WT-AZTMP	780 ± 64	9100 ± 540	21 ± 0.17	0.58
F105Y-TMP	430 ± 49	9000 ± 410	19 ± 0.17	1.2
F105Y-AZTMP	610 ± 59	10000 ± 480	19 ± 0.16	5.5

Table 4.8: Exchange parameters for Val95γ2

Protein-Ligand	k_{ex} (s^{-1})	ϕ_{ex} ($rad^{-2} \cdot s^{-2}$)	R_2^0 (s^{-1})	χ^2/N
WT-TMP	590 ± 38	16000 ± 350	19 ± 0.16	6.0
WT-AZTMP	1200 ± 72	17000 ± 820	19 ± 0.20	1.4
F105Y-TMP	500 ± 38	11000 ± 440	20 ± 0.15	1.5
F105Y-AZTMP	590 ± 57	13000 ± 470	19 ± 0.17	5.1

Table 4.9: Exchange parameters for Val10γ2

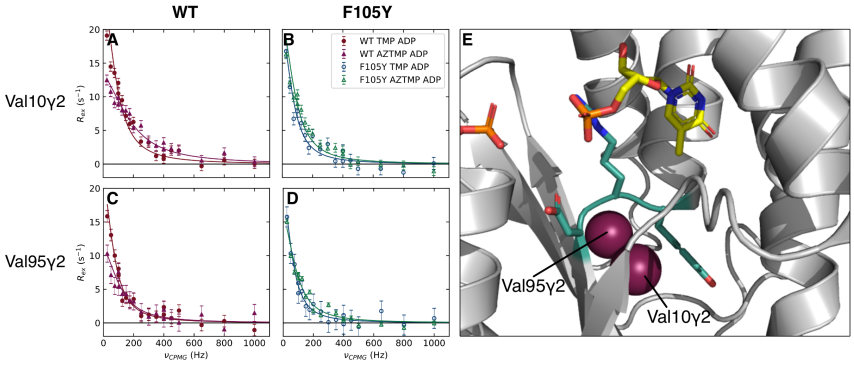


Figure 4.11: The active site dynamics of WT hTMPK are perturbed by AZTMP. (A-D) 850 MHz dispersion curves for WT complexes with TMP ADP (circles) and AZTMP ADP (triangles). F105Y complexes with TMP ADP (circles) and AZTMP ADP (triangles). R_2^0 is subtracted from the data, normalizing the curve to 0 to facilitate comparisons. (E) Val10γ2 and Val95γ2 are directly beneath the catalytically important DRY (teal sticks) motif. The F105Y-AZTMP and WT-TMP curves and parameters the most similar, possibly explaining why both have higher catalytic rates.

complexes have k_{ex} more similar to WT-TMP, including Ap_5dT which should closely approximate the transition state (Ostermann et al., 2000b; Feldhaus et al., 1975). We take this to suggest that AZTMP has a deleterious effect on the dynamics of Arg97 in the DRX motif, possibly by preventing the 2.1 Å swing of Asp15 into the NMP binding site. The increased exchange rates may represent Arg97 moving unproductively in the WT-AZTMP complex, while in the other complexes it is partially held in place by the availability of surrounding interacting partners. The F105Y complexes appear to be the more rigid in the ATP-binding site allowing Asp15 to interact with Arg97, stabilizing the active site by assuming the excited closed state.

If P-loop closure alone was the rate limiting step in catalysis than we would expect an increase in k_{ex} for Val14γ1 in F105Y-AZTMP versus WT-AZTMP that is consistent with the 20-fold difference in their catalytic rates. However, a mere 1.5-fold increase is observed. This, as well as other more widespread dynamical perturbations suggest a more nuanced mechanism. This leads us to suggest a tuning model for efficient phosphorylation of AZTMP by F105Y hTMPK. AZTMP is destabilizing to the active site of the enzyme, where the F105Y mutation is overly stabilizing leading to a decrease in catalytic rate with TMP. Both of these effects combined counterbalance the instability caused by AZTMP with conformational exchange occurring on approximately the same timescale in F105Y-AZTMP as WT-TMP.

Not all regions of the protein have their dynamics affected by the mutation or different sets of ligands. Leu167δ1 is located at the base of α6 and its exchange rates are similar for all four complexes, suggesting it is monitoring a motion that is unrelated to the catalytic cycle. A number of residues do not seem to show conformational exchange in any of the four (WT or F105Y with TMP or AZTMP) complexes discussed here. These residues are mostly located in the hydrophobic core, distal from the active site (Figure 4.12). Interestingly,

Val94γ2 which resides on the opposite side of the β-sheet from Val95γ2 does not show exchange despite being so close in sequence. This suggests that dynamical perturbations are local to the ligand binding and catalytic domains and reinforces the fact that the average solution structures of the complexes studied here are the same.

4.6 Conclusions

Despite minor structural differences between WT hTMPK bound to either TMP or AZTMP and ADP from crystallography and 2D NMR spectra, it is clear that these complexes show dynamical disparity on the μs-ms time scale. CPMG relaxation dispersion techniques reveal that AZTMP induces widespread dynamical perturbations to the P-loop, LID and at the catalytically important DRX motif. These perturbations likely stem from, but are not limited to, the steric clash between the 3'-azido group of AZTMP and Asp15 preventing its formation of hydrogen bonds with Arg97 of the DRX motif and also with Gln157 on α7. The F105Y mutation is able to overcome this steric clash by effectively pinning the P-loop to the NMP-binding site. This reduces conformational dynamics in the ATP-binding site resulting in increased affinity for the product ADP. The high activity of F105Y hTMPK towards AZTMP stems from proper tuning of the dynamics of the active site in the AZTMP complex, which mimics the WT-TMP complex and predisposes Arg45 and Arg97 into the catalytically active positions. Though, in the F105Y Ap_5dT complex a global reduction in the excited state may explain the reduced activity for TMP. We suggest intrinsic motions of the protein play a key role in catalysis and contribute to the differences in the catalytic efficiency between the different protein-ligand complexes. These insights into the intricate internal motions which dictate substrate specificity may serve as a guide for future designs of

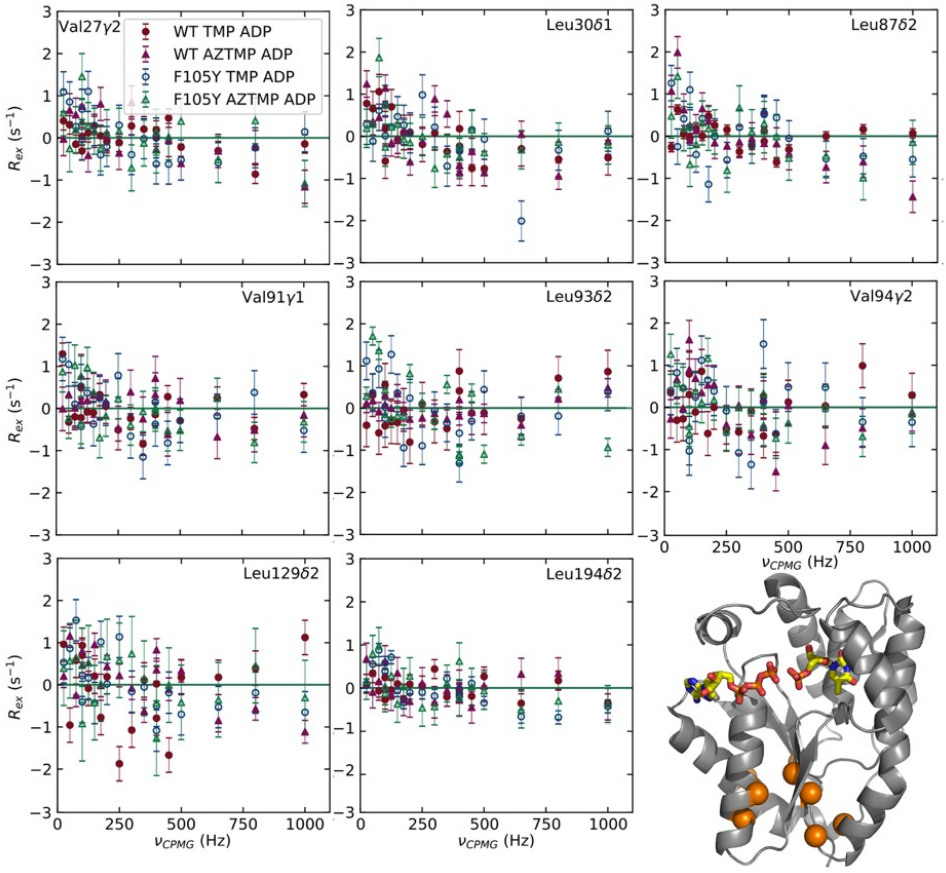


Figure 4.12: Residues showing no exchange in WT/F105Y-TMP/AZTMP complexes. These residues (orange spheres) map to the core of the protein, away from the active site, suggesting dynamical perturbations are local to the ligand binding region.

allosteric activators of hTMPK that will be useful to enhance the activity of AZT or inhibitors that would be useful in cancer chemotherapy. A better understanding of the human enzyme will also assist in the development of agents which are specific for anti-malarial (Chen et al., 2018) and anti-fungal (Huang et al., 2019) thymidylate kinases.

Chapter 5

`frenchdip`: a Python Program for Fitting CPMG Dispersion Curves

CPMG relaxation dispersion is one of many NMR techniques used for studying chemical exchange phenomena. The CPMG experiment can detect molecular motions on the μs-ms time scale, granted they occur at a rate of approximately 100 - 5000 s^{-1} (Palmer and Koss, 2019). This time scale is particularly relevant in the study of biomacromolecules as many important processes including folding, catalysis and molecular recognition occur at these rates. Not only is the interpretation of the results of relaxation dispersion difficult, but data analysis is a considerable roadblock for novices. This is not because a lack of software options, at least 8 are available freely from the internet or upon request (Hansen et al., 2008b; Bieri and Gooley, 2011; Kleckner and Foster, 2012; Sugase et al., 2013). Many dispersion curve fitting programs have even adopted the open source model, meaning that not only is the program free to download but also the source code can be modified and reused to fit the

users needs. Unfortunately, many of these programs lack documentation or a sufficiently intuitive interface leading to further confusion for novices (Kleckner and Foster, 2012).

Luckily in the age of "big data," non-linear curve fitting is quite trivial using freely available packages for many programming languages (Virtanen et al., 2019). For this reason, many laboratories choose to implement their own software. Certain aspects of CPMG data make analysis particularly challenging, including highly correlated parameters and relatively large datasets. Additionally, in order to obtain accurate fits for exchange parameters data must be collected at multiple static magnetic fields (Millet et al., 2000; Kovrigin et al., 2006). This requires multiple datasets to be fit in a global fashion, adding another layer of complexity to implementation of a fitting routine. Many labs also acquire data at multiple temperatures, ligand concentrations, pressures, etc. resulting in higher dimensional data, which needs to be handled appropriately. Available fitting software addresses many of these needs, however their complexity is often a deterrent. In Section 5.1 the advantages and missing features of available CPMG dispersion curve fitting software are discussed. In Section 5.2 the implementation of a new piece of dispersion curve fitting software, `frenchdip` [fitting relaxation (chemical exchange) dispersion in Python], is described. Finally, future prospects for `frenchdip` are discussed, focusing on implementation of currently missing features and its eventual release as well-documented open source software.

5.1 Comparison of Available CPMG Fitting Software

All CPMG data fitting software performs essentially the same task: reading of data, fitting data using some type of minimization routine and error estimation. Many pieces of fitting software (`GLOVE`, `CPMGfit`) combine configuration

files and data storage, which requires the user to generate multiple files for fits to different models (Sugase et al., 2013). Other programs (GUARDD, NESSY) require data to be properly formatted, then residues to be individually selected for fitting in a graphical interface (Kleckner and Foster, 2012; Bieri and Gooley, 2011). The program relax allows data to be read and converted into file formats compatible with most CPMG data fitting software (Bieri et al., 2011). Fragmentation in data formatting serves as another major roadblock for data analysis for the novices. Also of note is how each program defines the independent variable. Some choose to use the rate of 180° pulses in the CPMG pulse train (ν_{CPMG}) while others use the reciprocal of length of the delay between two 180° pulses ($1/\tau_{CP}$). An issue arrises because some groups define ν_{CPMG} in terms of the half delay between two 180° pulses ($1/\tau_{CPMG}$, n.b. the subscript of CPMG not CP) as $\nu_{CPMG} = 1/(4\tau_{CPMG})$. Thus the proper conversion is $1/\tau_{CP} = 2\nu_{CPMG}$ and ν_{CPMG} can be defined as $n_{cyc}/\mathrm{T}_{relax}$ where n_{cyc} is the number of CPMG elements and T_{relax} is the total time of the relaxation delay (Sugase et al., 2013). This confusing notation may lead to a factor of 2 error for which the user is entirely responsible.

Dispersion curve fitting software usually differs in programming language and minimization technique. Many older programs are implemented in FORTRAN, while newer ones are implemented in Python and MATLAB. The choice of programming language does not matter a great deal, however those programs that use MATLAB require an expensive license to use. Fortunately, many institutions provide access to MATLAB, though this does limit the use of fitting software written in this language. Most fitting software is based on using grid search to minimize the χ^2 over a large range of values for each parameter. Some programs include additional minimization techniques in an effort to reduce fitting time, however these methods do not always converge on the global minimum. Experimental data are fit to models (reproduced from Chapter 3)

of chemical exchange suitable for all exchange regimes (Carver and Richards, 1972):

$$
\begin{aligned}
R_{2,eff} &= R_2^0 + \frac{k_{ex}}{2} - \nu_{CPMG}\cosh^{-1}[D_+\cosh(\eta_+) - D_- cos(\eta_-)] \\
D_\pm &= \frac{1}{2}(\pm 1 + \frac{\Psi + 2\Delta\omega^2}{\sqrt{\Psi^2+\zeta^2}}) \\
\eta_\pm &= \frac{1}{2\sqrt{2}\nu_{CPMG}}\sqrt{\pm\Psi + \sqrt{\Psi^2+\zeta^2}} \\
\Psi &= (\Delta R_2^0 + p_a k_{ex} + p_b k_{ex})^2 - \Delta\omega^2 + 4p_a p_b k_{ex}^2 \\
\zeta &= 2\Delta\omega(\Delta R_2^0 + p_a k_{ex} + p_b k_{ex})
\end{aligned}
\tag{5.1}
$$

or to one suitable for fast exchange (Luz and Meiboom, 1963):

$$
R_{2,eff} = R_2^0 + \frac{\phi_{ex}}{k_{ex}}[1 - \frac{\nu_{CPMG}}{k_{ex}}\tanh(\frac{k_{ex}}{4\nu_{CPMG}})] \tag{5.2}
$$

where $\phi_{ex} = p_a p_b \Delta\omega^2$. Other models exist for N-state systems, but the two-state models in Equation 5.1 and 5.2 are usually sufficient (Koss et al., 2018). Error estimation is performed using a Monte Carlo or jackknife method, though implementation specifics differ between software.

A comparison of two freely available CPMG dispersion curve fitting software and `frenchdip` is presented in Table 5.1. All programs fit similar values for each exchange parameter when fit to Equation 5.1, but differ slightly in their estimate of the error. Interestingly, `GUARDD` does not implement Equation 5.2, and can only fit to the full equation. This is problematic when fitting data from residues that are in fast exchange, as was the case for the ILV methyls from hTMPK. Globally fitting multiple residues to single k_{ex} or p_a but not $\Delta\omega$ is not possible in any available software.

5.2 Design and Implementation of `frenchdip`

`frenchdip` is implemented in the Python programming language, using the NumPy (van der Walt et al., 2011) library for data storage and the SciPy

Software	k_{ex} (s^{-1})	$\Delta\omega$ (ppm)	p_a (%)	R_2^0 (s^{-1})
CPMGfit	2700 ± 80	0.62 ± 0.016	98.5 ± 1.9	13.1 ± 0.29
GUARDD	2100 ± 260	0.62 ± 0.030	98.6 ± 0.10	15.2 ± 1.0
frenchdip	2200 ± 360	0.62 ± 0.014	98.6 ± 0.080	14.7 ± 1.5

Table 5.1: Comparison of CPMG fitting software for the dispersion curve of WT hTMPK AZTMP 2 mM AZTMP Ile184δ1. All programs fit similar values for each parameter for a slow exchange dispersion curve. Values for R_2^0 and k_{ex} vary somewhat between programs.

library for data fitting (Virtanen et al., 2019). These two libraries, alone, make it quite simple to implement a data fitter in very few lines of code. The data was fit by a grid search method (scipy.optimize.brute) and model selection performed by comparing values of reduced χ^2. Fitting errors can be estimated using either a Monte Carlo or bootstrap method. Bootstrap error estimation is appropriate in cases where the number of data points is large, and the curve is appropriately sampled. In instances where this is not true the errors will be very large, however, this does allow the user to diagnose issues of improper sampling in their data.

The design of `frenchdip` seeks to address two missing features of other CPMG data fitting software: (1) better batch fitting methods and (2) explicit global fitting. Batch fitting refers to successive fitting of datasets without user interaction. `GUARDD`, perhaps the most user-friendly piece of CPMG fitting software, includes a batch mode, however significant user interaction is required in order to queue multiple fits. This is solved in `frenchdip` by allowing users to define a model or set of models that data should be fit to as well as the ability to group datasets together and choose which parameters should be shared and which should be global. Many command line based tools require user input for each individual fit, which requires a separate script to be written by the user to automate the process. In the context of CPMG data fitting the term global fitting is routinely used to describe using data collected at multiple fields to a single set of exchange parameters and scaling of $\Delta\omega$ by the ratio of the static fields. Here, the term "global fitting" or "global parameters" are used to describe a situation where dispersion curves from multiple residues are grouped and fit together. The term "shared parameters" is used to describe cases where a single dataset consisting of multiple curves from data collected at static fields, temperatures, etc. is fit to the same k_{ex}, $\Delta\omega$ and p_a and scaled appropriately.

Two aspects of the implementation allow for the above design goals to be

met. The first is data storage, at the time of writing, data is stored in a simple comma separated value (CSV) format, which is compatible with all major spreadsheet programs as well as simple text editors. By separating the data file from a settings file, unlike `CPMGfit` and `GLOVE`, datasets can be fit to multiple models with minimal user interaction. This also allows multiple residues' CSVs to be read and combined for global fitting routines without requiring a combined dataset file to be created by the user.

The second, and most important implementation detail is the use of indexer arrays to simplify the creation of objective functions with shared and/or global parameters. The fitting routines found in the scipy.optimize submodule accept a 1-dimensional array of parameters to be minimized. In order to specify an objective function with shared parameters, this array must contain a single copy of each shared parameter and as many copies as there are datasets for other parameters. For the simple case, assume that a single residue has two dispersion curves collected at two different fields and are to be fit to the fast exchange model (Equation 5.2). This means the parameters k_{ex} and ϕ_{ex} should be shared and two different values for R_2^0 should be fit. The parameter array would look like this:

$$p = [k_{ex}, \phi_{ex}, R^0_{2,1}, R^0_{2,2}]$$

To define a curve for each data set this parameter array must be properly indexed to produce the objective function:

$$\begin{bmatrix} data0 \\ data1 \end{bmatrix} - \begin{bmatrix} FastExchange(k_{ex}, \phi_{ex}, R^0_{2,1}) \\ FastExchange(k_{ex}, \phi_{ex}, R^0_{2,2}) \end{bmatrix}$$

It is much simpler to pass an array of parameters to the *FastExchange* function. A unique quality of NumPy arrays is that they can be indexed by other NumPy arrays, or indexer arrays (van der Walt et al., 2011). The indexer array i and

indexed parameter array $p[i]$,

$$i = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 1 & 3 \end{bmatrix}, \; p[i] = \begin{bmatrix} k_{ex} & \phi_{ex} & R^0_{2,1} \\ k_{ex} & \phi_{ex} & R^0_{2,2} \end{bmatrix}$$

produce the desired outcome, such that the objective function is:

$$data - FastExchange(p[i])$$

where square brackets indicate array indexing. Using an example with both shared and global parameters it will be demonstrated that this is a general solution to programmatically defining the objective function. In the second example 3 residues will be fit together, each with a curve collected at two static fields (6 total). The curves will again be fit the fast exchange model where ϕ_{ex} is a shared parameter. In this case k_{ex} will be a global parameter. Now, the parameter array would look like this:

$$p = [k_{ex}, \phi_{ex,1}, R^0_{2,1}, R^0_{2,2}, \phi_{ex,2}, R^0_{2,3}, R^0_{2,4}, \phi_{ex,3}, R^0_{2,5}, R^0_{2,6}]$$

the indexer and indexed parameter arrays like this:

$$i = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 4 & 5 \\ 0 & 4 & 6 \\ 0 & 7 & 8 \\ 0 & 7 & 9 \end{bmatrix}, \; p[i] = \begin{bmatrix} k_{ex} & \phi_{ex,1} & R^0_{2,1} \\ k_{ex} & \phi_{ex,1} & R^0_{2,2} \\ k_{ex} & \phi_{ex,2} & R^0_{2,3} \\ k_{ex} & \phi_{ex,2} & R^0_{2,4} \\ k_{ex} & \phi_{ex,3} & R^0_{2,5} \\ k_{ex} & \phi_{ex,3} & R^0_{2,6} \end{bmatrix}$$

and the objective function remains:

$$data - FastExchange(p[i])$$

highlighting the advantage of passing an array of parameters to the model function. Because these arrays are generated programmatically, the user only needs

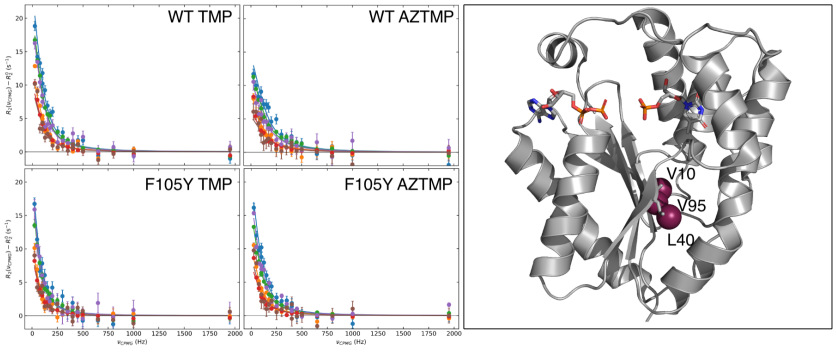


Figure 5.1: Globally fit dispersion curves for V10, L40 and V95. The dispersion curves for these three residues were fit to a single k_{ex} in as described by the second example in Section 5.2. The positions of these residues on the protein is shown (right) with each methyl group shown in red spheres.

Protein·Ligands	k_{ex} (s^{-1})	$\phi_{ex,V10\gamma2}$ ($rad^{-2}\cdot s^{-2}$)	$\phi_{ex,L40\delta1}$ ($rad^{-2}\cdot s^{-2}$)	$\phi_{ex,V95\gamma2}$ ($rad^{-2}\cdot s^{-2}$)
WT·TMP·ADP	650 ± 38	16000 ± 820	13000 ± 490	12000 ± 680
WT·pAZT·ADP	940 ± 78	14000 ± 1200	1200 ± 779	10000 ± 880
F105Y·TMP·ADP	500 ± 39	1100 ± 800	9000 ± 482	9400 ± 550
F105Y·pAZT·ADP	650 ± 47	13000 ± 810	9900 ± 488	11000 ± 640

Table 5.2: Global fitting of k_{ex} to 3 nearby residues in hTMPK

to choose which parameters are shared and which are global. This method was used to globally fit three residues in hTMPK to the fast exchange model (Figure 5.1, Table 5.2). The global fit to k_{ex} is reasonable for this cluster of residues, suggesting their motions might be coupled.

5.3 Future Perspectives

A number of features for `frenchdip` are unfinished. Currently, a single script exists which reads all properly formatted CSV files in a directory and fits them to the full model (Equation 5.1), fast exchange model (Equation 5.2) and a no exchange model. A second, custom purpose script was used to do the global fitting in Figure 5.1. Handling of the parameter arrays and model functions has since been rewritten using an object-oriented approach. This allows individual parameters to be set to shared or global, and enables joint fitting of residues without requiring a custom purpose script. Remaining features to be implemented include: fitting data collected at multiple temperatures, or other physical quantities and scaling shared parameters accordingly, a statistical method for model selection (Akaike and/or Bayesian Information Criterion), a command line interface, standardizing an input file format. Model selection is particularly important as it would serve as an estimator how appropriate global fitting is cases where k_{ex} and p_b are similar for a subset of residues (Kumar et al., 2018). Finally, once fully featured, releasing `frenchdip` as an open-source project so that it would be available to the wider community.

Chapter 6

Materials and Methods

Protein Expression and Purification

Rho130 was expressed from pET11a as previously described (Briercheck et al., 1996). The synthetic gene codon-optimized for expression of the N-terminus (His)6 tagged hTMPK in *E. coli* was obtained from ATUM and cloned into the pET22b(+) vector using the NdeI and XhoI sites. All proteins were expressed in C3013 T7 Express *lysY/lacIq E. coli* cells (New England Biolabs). Unlabeled samples were grown in Luria-Bertani broth containing 100 mg/mL ampicillin at 30 °C to an OD_{600} of 0.8 and induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Uniformly ^{15}N-labeled protein was produced in Studier's PG medium (Studier, 2005) with $^{15}NH_4Cl$ as the sole nitrogen source. Uniformly ^{13}C-labeled protein were grown in the same media with D-glucose-$^{13}C_6$ as the sole carbon source. Specifically lysine labeled protein was produced in ^{14}N PG with [U-^{13}C, U-^{15}N] lysine added 1 h prior to induction with IPTG. Specifically labeled $^{13}CH_3$ and $^{13}CHD_2$ ILV labeled protein was grown in ^{15}N PG with $^{13}CH_3$ or $^{13}CHD_2$ α-ketoisovalerate (100 mg/L) and α-ketobutyrate (50 mg/L) added 1 h prior to induction for 20 h for hTMPK and 9 h for Rho130 at an OD_{600} of 1.6 (Goto et al., 1999). For Rho130 guanosine,

cytidine and uridine (50 mg/L each) were also added at the time of induction. Cells were stored at -80 °C after harvesting.

For Rho130, Lysates were loaded onto a CM-sephadex column, washed and eluted with CM buffer (50 mM Tris-HCl, 10 mM EDTA, pH 7.6) and a 0-0.8 M NaCl gradient, concentrated fractions were loaded onto a G50 sephadex gel filtration column and eluted using G50 buffer (10 mM K_2PO_4, 150 mM K_2SO_4, 5 mM EDTA, pH 7.0). Fractions containing Rho130 were concentrated, exchange into DEAE buffer (50 mM Tris-HCl, 10 mM EDTA, pH 8.0) and loaded onto a DEAE-sephadex column and eluted with a 0-0.3 M K_2SO_4 gradient, concentrated and dialyzed into NMR buffer (G50 buffer + 0.02% NaN_3) or PBPE reaction buffer (50 mM K_2PO_4, 150 mM K_2SO_4, 0.02% NaN_3, pH 7.8), pH adjustments were performed using H_3PO_4 so as not to include excess halide salts. Purity was assayed by SDS-PAGE and photometrically, with an A275/A290 ¿ 6 deemed acceptable.

For hTMPK, The cell pellets were resuspended in lysis buffer (50 mM sodium phosphate, 300 mM sodium chloride, 10 mM imidazole, 20 mM $MgCl_2$, 10% (w/v) glycerol, 50 μM TMP, pH 7.4) and lysed by addition of 5 mg lysozyme and sonification. The lysate was centrifuged at 20,000g for 30 min at 4 °C. HisPur Cobalt Resin (Thermo Scientific) added to the supernatant and mixed on an end-over-end rotator for 1 h. This supernatant/resin mix was then poured into a gravity chromatography column to separate the resin from the flow-through. The resin was then washed with 3 column volumes of wash buffer (50 mM sodium phosphate, 300 mM sodium chloride, 10 mM imidazole, pH 7.4). The protein was eluted with the elution buffer (50 mM sodium phosphate, 300 mM sodium chloride, and 500 mM imidazole, pH 7.4). The protein was then dialyzed against the cation exchange buffer (50 mM sodium phosphate, 50 mM sodium chloride, 10 mM EDTA, pH 7.0) and loaded on a SP Sephadex C-50 column equilibrated with the same buffer and eluted with

a 0.05 to 1.5 M NaCl gradient and concentrated. hTMPK was then dialyzed into NMR buffer (20 mM MES, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, 0.02% (w/v) NaN3, pH 6.9)

Rho130-Initiator Attachment

ATRP initiator immobilization was performed as previously reported (Murata et al., 2013). To a stirring solution of Rho130 (8.7 mg/mL) in PBPE reaction buffer was added 50x molar excess N-2-bromo-2-methylpropionyl-β-alanine N'-oxysuccinimide ester (NHS-initiator). The NHS-initiator was dissolved in a minimal volume (final concentration 10% (v/v)) DMF before addition. Rho130-initiator conjugates were dialyzed in a 3K MWCO spin concentrator (Pall Corp.) after 3 h. Initiator conjugates were used as is with no further purification.

Grafting-from Rho130-Initiator Conjugates

Rho130-pCAm and Rho130-pQA conjugates were synthesized by Chad Cummings, Hironobu Murata and Bibifatima Kaupbayeva of the Russell laboratory at Carnegie Mellon University. To a 0.11 mM (1 mM initiator in the 9-initiator conjugate) solution of Rho130-initiator conjugates and a 10 equiv of 6-(methacryloylamino)hexanoic acid (CAm) were added to a round bottom flask and bubbled with dry argon for 30 min. To a separate flask were added Cu(II)Cl (10 equiv), sodium ascorbate (1 equiv) and 1,1,4,7,10,10-hexamethyltriethylenetetramine (HMTETA) and bubbled with dry argon for 10 min. The copper/ligand complex was added to the Rho130-initiator/monomer mixture and stirred for 1 h. The reaction was stopped by opening to the air.

Gel Electrophoresis

Stacking gels were 4% acrylamide in 250 mM Tris-HCl (pH 6.8) and separating gels were 12% acrylamide in 250 mM Tris-HCl (pH 8.8). Gels were polymerized with 25 μL TEMED and 25 μL 10% ammonium persulfate solution. Samples were run at 110 V until the loading dye reached the end of the gel. Gels were visualized with Coomassie blue.

^{15}N Resonance Assignment in Rho130-initiator conjugates

NMR spectra were acquired on a 600 MHz Bruker Avance AMX spectrometer with a room temperature probe. Samples were prepared in NMR buffer. ^{15}N HSQC, ^{15}N-NOESY-HSQC and ^{15}N-HSQC-TOCSY experiments were performed on ^{15}N-labeled samples, HNCA, HNCACB and HCCH-TOCSY experiments were performed on uniformly ^{13}C-labeled samples. Data was collected for both unmodified and initiator-modified protein for the ^{15}N HSQC, while the ^{15}N-HSQC-TOCSY, HNCA, HNCACB and HCCH-TOCSY were performed on only initiator-modified protein.

Proteolysis

Digestions were performed with trypsin, LysC and GluC. In each case Rho130 or Rho130-initiator conjugates were denatured with 6 M guanidine hydrochloride, 50 mM Tris, pH 8.0 at 37 °C. Samples were exchanged into 50 mM Tris, pH 8.0 before addition of the protease. Proteases were added in a 1:20 (w/w) to Rho130 and Rho130-initiator conjugates. The reaction was performed for 18 h at 37 °C after which samples were precipitated with an equal volume of 20% trichloroacetic acid, pelleted, washed with cold acetone and reconstituted at 50μM in MS buffer (50/50 acetonitrile/H_2O, 0.1 % acetic acid).

Mass Spectrometry

Mass spectrometry (MS) was performed on a Thermo-Fisher LCQ Fleet ESI Ion Trap. ^{15}N-labeled samples were prepared in MS buffer at 50 μM. Samples of Rho130, Rho130-initiator conjugates, Rho130-pCAm conjugates, protease-treated unmodified Rho130 and protease-treated Rho130-initiator conjugates were analyzed. Data were processed and analyzed using the Xcalibur software. Data were averaged over the time immediately following injection to produce the final spectrum. Peaks were assigned manually, by calculating the m/z for all possible initiator conjugates and peptides. The estimated molecular weight was divided by peak position calculate the charge.

^{13}C ILV Methyl Spectroscopy

^{13}C HMQC spectra were acquired on a 600 MHz Bruker AMX spectrometer equipped with a room temperature probe, 4D methyl HMQC-NOE-HMQC spectra were acquired on a 700 MHz Bruker-BioSpin Avance spectrometer equipped with a cryo-probe, CHD_2 (hTMPK) and CH_3 (Rho130) methyl 1H Carr-Purcell-Meiboom-Gill (CPMG) experiments (Baldwin et al., 2010) and CHD_2 methyl ^{13}C R_1 and $R_{1\rho}$ (Tugarinov and Kay, 2005) experiments were acquired on a 600 MHz and 850 MHz Bruker AVIII spectrometers both equipped with cryo-probes. All samples were prepared as 0.4-0.8 mM solutions of protein in their respective NMR buffer dissolved in 100% D_2O. For hTMPK, TMP, AZTMP and ADP at a concentration of 2 mM were added for HMQC, NOE, and R_1 and $R_{1\rho}$ experiments. For the CPMG experiments the concentration of ADP was raised to 8 mM to fully saturate the enzyme. A concentration of 0.84 mM Ap_5dT was used for CPMG experiments. Spectra were taken at 34 °C. CHD_2 methyl ^{1}H CPMG were acquired with a T_{relax} (total time of the CPMG pulse train) of 40 ms and taken for n_{cyc} (number of CPMG elements) of 1, 2,

3, 4, 4, 5, 6, 7, 8, 10, 12, 14, 16, 16, 18, 20, 26, 32, 40 and 78. CPMG data for TMP and AZTMP samples were acquired at 600 MHz and 850 MHz. CPMG data for Ap_5dT was acquired at 850 MHz. All spectra were processed with NMRPipe and visualized with nmrDraw (Delaglio et al., 1995). Dispersion curves were produced by calculating $R_{2,eff}(\nu_{CPMG}) = -\frac{1}{T_{relax}} ln \frac{I(\nu_{CPMG})}{I(0)}$ and plotting against ν_{CPMG} data was fit using `frenchdip` (Chapter 5)

ILV Methyl Resonance Assignments in hTMPK

Resonance assignment of hTMPK was performed by Kaustubh Sinha. Ile (δ1), and stereospecific Leu (δ1 and δ2), and Val (γ1 and γ2) assignments of the WT-TMP/ADP/Mg^{2+} complex were obtained by a combination of site-directed mutagenesis, pro-S labeling of LV methyl groups, four-dimensional (4D) and three-dimensional (3D) methyl–methyl NOE. The availability of Val and Leu assignments provided greater coverage of the enzyme. Single-site (Ile $\rightarrow$ Val) mutations were generated for all the Ile residues and a few selected Leu and Val residues (Val $\rightarrow$ Ala and Leu $\rightarrow$ Val). The ^{13}C HMQC spectra of the mutant proteins that showed a single missing resonance (Ile mutants) or a pair of missing resonances (Val or Leu mutants) with minor movement of the remaining ILV peaks were used for assignments. NMR spectra of all of the Ile mutants were essentially identical to the WT spectrum, with the exception of a missing peak corresponding to the altered Ile residue. Specific labeling of the pro-S methyl groups of Leu (δ2-CH_3) and Val (γ2-CH_3) was achieved by expressing the protein in the presence of 2-[$^{13}CH_3$]methyl-4-[$^2H_3$]acetolactate (NMR-Bio) (Gans et al., 2010). The resulting methyl spectra consists of only the pro-S methyl peaks. Peaks that were not labeled using this protocol were identified as pro-R peaks. The X-ray structures of the human thymidylate kinase in complex with TMP and ADP (PDB 1E2G) permitted the use of NOE experiments to assign additional ILV methyl peaks and confirm the peaks already

assigned. The 4D methyl HMQC-NOE-HMQC pulse program was obtained from Clore (http://spin.niddk.nih.gov/clore/Software/software.html) (Vuister et al., 1993). Two differently labeled ILV [^{13}C]methyl-labeled samples were produced for the 4D NOE experiments. The first NOE sample was produced using 4-[^{13}C]-α-ketobutyric acid (2-KB) and dimethyl[$^{13}C_2$]-α-ketoisovaleric acid as the precursors for ILV labeling (Tugarinov and Kay, 2003). The precursor dimethyl[$^{13}C_2$]-α-ketoisovaleric acid introduces ^{13}C into all LV methyl groups and gives a strong intra-residue NOE peak in the 4D NOE spectrum. This allowed identification of the methyl pairs for LV residues. The second sample was produced from precursors 2-KB and 3-methyl[^{13}C]-3,4,4,4-[$^{2}H_4$]-α-ketoisovaleric acid. This resulted in LV side chains with $^{13}CH_3$ and $^{12}CD_3$ methyl groups. The NOE spectrum of this sample lacked the strong intra-residue peak and permitted detection of NOEs over longer distances due to the absence of transfer to the other methyl on the same residue. 4D NOE data at one mixing time along with the mutagenesis and specific labeling experiments were not enough to confidently assign some of the residues. In such cases, we relied on the dependence of NOE intensities on mixing time. A series of 3D methyl NOE experiments were acquired with different NOE mixing times (100ms to 900 ms) and the experimental data was compared to the simulated data. This approach helped in resolving inconclusive assignments and re-affirmed previous assignments.

Measuring ADP Binding Affinity

A series of ^{13}C HMQC spectra were acquired for concentrations of ADP ranging from 0-2.4 ligand-to-protein ratio for WT or F105Y bound to either TMP or AZTMP. Spectra were processed as above, except exponential line broadening was used instead of sine-bell apodization. The observed dissociation constant (K_D) and off-rate (k_{off}) for ADP with either TMP or AZTMP were deter-

mined using two-dimensional line fitting using the TITAN program (Waudby et al., 2016), utilizing five resonance lines which moved upon addition of ADP (Leu161δ1, Ile184δ1, Ile191δ1, Ile199δ1, Val193δ1) for each of the four samples. The simple 2-state binding model [P] + [L] $\rightleftharpoons$ [PL] was chosen, and errors were estimated using the built-in bootstrap procedure in TITAN.

^{13}C Spin Relaxation Data Analysis

The ^{13}C R_1 and $R_{1\rho}$ relaxation rates were used to extract S^2 order parameters for each of the ILV methyl groups using the method of Tugarinov and Kay (Tugarinov and Kay, 2005). The anisotropic rotational diffusion tensor used in the calculations was obtained using HYDRONMR (Garcia de la Torre et al., 2000) and the X-ray structure of hTMPK (PDB: 1E2D). The local proton concentration was calculated using an in-house R script. S^2 order parameters were calculated using the ^{13}C relaxation measurements, diffusion tensor and proton concentration using MATLAB scripts provided by Tugarinov.

Appendix A

Glossary

chemical shift – resonance frequency of a nucleus relative to a reference standard. In an NMR spectrum it is the location of the peak on the spectrum.

chemical shift perturbation (CSP) – when a treatment, such as ligand binding, causes the position of an atom's peak in the NMR spectrum to move. Indicates a change in the environment around that atom.

CPMG relaxation dispersion – Carr-Purcell-Meiboom-Gill relaxation dispersion – an NMR experiment that assays chemical (or conformational) exchange. This is achieved by detecting the R_2 relaxation rate as a function of the rate of refocusing pulses. The refocusing pulses effectively remove the exchange component to R_2. The resultant dispersion curves (R_2 vs. pulse rate) are fit to models of chemical exchange to extract information about the process (k_{ex}, $\Delta\omega$, p_a).

HSQC – heteronuclear single quantum coherence – a two-dimensional NMR experiment with one axis being ^{1}H and the other the heteronucleus (^{13}C or ^{15}N). A single peak appears for each unique proton connected to the heteroatom.

HNCA/HNCACB – triple-resonance experiment used for the assignment of amide backbone peaks in the ^{15}N HSQC. These three-dimensional experiment are named for the magnetization transfer pathway ($H_N \rightarrow N \rightarrow C\alpha$, and additionally to $C\beta$ for the HNCACB). The resulting spectrum shows two peaks in the ^{13}C dimension for each peak in the ^{15}N HSQC corresponding to the i peak or self peak, and i-1 peak corresponding to the $C\alpha$ from the preceding residue.

HMQC – heteronuclear multiple quantum coherence – a similar experiment to an HSQC but performed in a different way. This experiment has advantages for detecting methyl groups in large proteins.

NOESY – nuclear Overhauser effect spectroscopy – a two- three- or four-dimensional NMR experiment in which peaks arise from proximity to one another. Nuclei are irradiated and allowed to mix for a set amount of time, this allows population differences to be transferred to nearby atoms. A spectrum is acquired after mixing and signal intensities from the other atoms are proportional to $\frac{1}{d^6}$ where d is the distance between the atoms.

Atom Nomenclature. Atoms are referred to by the name and position of the residue. In cases where stereospecific assignments have been acquired atoms are referred to by their three letter residue code, residue index and atom name. Atoms are named using Greek letters, starting with $C\alpha$, the carbon in to the protein backbone. For instance Leu137δ1 refers to the $C\delta$ atom, or the carbon of one of the two methyl groups at the end of the amino acid side chain.

References

A. Abuchowski, T. van Es, N. C. Palczuk, and F. F. Davis. Alteration of immunological properties of bovine serum albumin by covalent attachment of polyethylene glycol. *The Journal of Biological Chemistry*, 252(11):3578–3581, June 1977. ISSN 0021-9258.

T. J. Allison, T. C. Wood, D. M. Briercheck, F. Rastinejad, J. P. Richardson, and G. S. Rule. Crystal structure of the RNA-binding domain from transcription termination factor rho. *Nature Structural Biology*, 5(5):352–356, May 1998. ISSN 1072-8368. doi: 10.1038/nsb0598-352. URL `http://www.nature.com/doifinder/10.1038/nsb0598-352`.

S. Averick, A. Simakova, S. Park, D. Konkolewicz, A. J. D. Magenau, R. A. Mehl, and K. Matyjaszewski. ATRP under Biologically Relevant Conditions: Grafting from a Protein. *ACS Macro Letters*, 1(1):6–10, 2012. doi: 10.1021/mz200020c. URL `http://pubs.acs.org/doi/abs/10.1021/mz200020c`.

S. L. Baker, A. Munasinghe, B. Kaupbayeva, N. Rebecca Kang, M. Certiat, H. Murata, K. Matyjaszewski, P. Lin, C. M. Colina, and A. J. Russell. Transforming protein-polymer conjugate purification by tuning protein solubility. *Nature Communications*, 10(1):4718, Dec. 2019a. ISSN 2041-1723. doi: 10.1038/s41467-019-12612-9. URL `http://www.nature.com/articles/s41467-019-12612-9`.

S. L. Baker, H. Murata, B. Kaupbayeva, A. Tasbolat, K. Matyjaszewski, and A. J. Russell. Charge-Preserving Atom Transfer Radical Polymerization Initiator Rescues the Lost Function of Negatively Charged Protein–Polymer Conjugates. *Biomacromolecules*, 20(6):2392–2405, June 2019b. ISSN 1525-7797, 1526-4602. doi: 10.1021/acs.biomac.9b00379. URL `https://pubs.acs.org/doi/10.1021/acs.biomac.9b00379`.

A. J. Baldwin, T. L. Religa, D. F. Hansen, G. Bouvignies, and L. E. Kay. ^{13}CHD$_2$ Methyl Group Probes of Millisecond Time Scale Exchange in Proteins by ^{1}H Relaxation Dispersion: An Application to Proteasome Gating Residue Dynamics. *Journal of the American Chemical Society*, 132(32):10992–10995, Aug. 2010. ISSN 0002-7863, 1520-5126. doi: 10.1021/ja104578n. URL `http://pubs.acs.org/doi/abs/10.1021/ja104578n`.

M. Bieri and P. R. Gooley. Automated NMR relaxation dispersion data analysis using NESSY. *BMC Bioinformatics*, 12(1):421, Dec. 2011. ISSN 1471-2105. doi: 10.1186/1471-2105-12-421. URL https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-421.

M. Bieri, E. J. d'Auvergne, and P. R. Gooley. relaxGUI: a new software for fast and simple NMR relaxation data analysis and calculation of ps-ns and µs motion of proteins. *Journal of Biomolecular NMR*, 50(2):147–155, June 2011. ISSN 0925-2738, 1573-5001. doi: 10.1007/s10858-011-9509-1. URL http://link.springer.com/10.1007/s10858-011-9509-1.

A. Biswas, A. Shukla, R. Vijayan, J. Jeyakanthan, and K. Sekar. Crystal structures of an archaeal thymidylate kinase from Sulfolobus tokodaii provide insights into the role of a conserved active site Arginine residue. *Journal of Structural Biology*, 197(3):236–249, Mar. 2017. ISSN 10478477. doi: 10.1016/j.jsb.2016.12.001. URL http://linkinghub.elsevier.com/retrieve/pii/S1047847716302477.

C. E. Bogden, D. Fass, N. Bergman, M. D. Nichols, and J. M. Berger. The structural basis for terminator recognition by the Rho transcription termination factor. *Molecular Cell*, 3(4):487–493, Apr. 1999. ISSN 1097-2765. doi: 10.1016/s1097-2765(00)80476-1.

Z. K. Boswell and M. P. Latham. Methyl-Based NMR Spectroscopy Methods for Uncovering Structural Dynamics in Large Proteins and Protein Complexes. *Biochemistry*, Oct. 2018. ISSN 0006-2960, 1520-4995. doi: 10.1021/acs.biochem.8b00953. URL http://pubs.acs.org/doi/10.1021/acs.biochem.8b00953.

D. M. Briercheck, T. J. Allison, J. P. Richardson, J. F. Ellena, T. C. Wood, and G. S. Rule. 1h, 15n and 13c resonance assignments and secondary structure determination of the RNA-binding domain of E.coli rho protein. *Journal of biomolecular NMR*, 8(4):429–444, 1996. URL http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=9008362&retmode=ref&cmd=prlinks.

D. M. Briercheck, T. C. Wood, T. J. Allison, J. P. Richardson, and G. S. Rule. The NMR structure of the RNA binding domain of E.coli rho factor suggests possible RNA–protein interactions. *Nature Structural Biology*, 5 (5):393–399, May 1998. ISSN 1072-8368. doi: 10.1038/nsb0598-393. URL http://www.nature.com/doifinder/10.1038/nsb0598-393.

R. Brundiers, A. Lavie, T. Veit, J. Reinstein, I. Schlichting, N. Ostermann, R. S. Goody, and M. Konrad. Modifying human thymidylate kinase to potentiate azidothymidine activation. *Journal of Biological Chemistry*, 274 (50):35289–35292, 1999. ISSN 0021-9258 (Print)\r0021-9258 (Linking). doi: 10.1074/jbc.274.50.35289.

C. Caillat, D. Topalis, L. A. Agrofoglio, S. Pochet, J. Balzarini, D. Deville-Bonne, and P. Meyer. Crystal structure of poxvirus thymidylate kinase: An unexpected dimerization has implications for antiviral therapy. *Proceedings of the National Academy of Sciences*, 105(44):16900–16905, 2008. ISSN 1091-6490 (Electronic)\r0027-8424 (Linking). doi: 10.1073/pnas.0804525105. URL `http://www.pnas.org/cgi/doi/10.1073/pnas.0804525105`.

S. Carmali, H. Murata, C. Cummings, K. Matyjaszewski, and A. J. Russell. Polymer-Based Protein Engineering. In *Methods in Enzymology*, volume 590, pages 347–380. Elsevier, 2017. ISBN 978-0-12-810502-3. doi: 10.1016/bs.mie.2016.12.005. URL `https://linkinghub.elsevier.com/retrieve/pii/S0076687916304335`.

J. Carver and R. Richards. A general two-site solution for the chemical exchange produced dependence of T2 upon the carr-Purcell pulse separation. *Journal of Magnetic Resonance (1969)*, 6(1):89–105, Jan. 1972. ISSN 00222364. doi: 10.1016/0022-2364(72)90090-X. URL `http://linkinghub.elsevier.com/retrieve/pii/002223647290090X`.

G. Cattani, L. Vogeley, and P. B. Crowley. Structure of a PEGylated protein reveals a highly porous double-helical assembly. *Nature Chemistry*, 7(10):823–828, 2015. ISSN 1755-4330. doi: 10.1038/nchem.2342. URL `http://dx.doi.org/10.1038/nchem.2342`.

S. K. Chaudhary, Y. Iyyappan, M. Elayappan, J. Jeyakanthan, and K. Sekar. Insights into product release dynamics through structural analyses of thymidylate kinase. *International Journal of Biological Macromolecules*, 123:637–647, Feb. 2019. ISSN 01418130. doi: 10.1016/j.ijbiomac.2018.11.025. URL `https://linkinghub.elsevier.com/retrieve/pii/S0141813018334007`.

M. D. Chen, K. Sinha, G. S. Rule, and D. H. Ly. Interaction of α-Thymidine Inhibitors with Thymidylate Kinase from *Plasmodium falciparum*. *Biochemistry*, 57(19):2868–2875, May 2018. ISSN 0006-2960, 1520-4995. doi: 10.1021/acs.biochem.8b00162. URL `http://pubs.acs.org/doi/10.1021/acs.biochem.8b00162`.

Y. H. Chen, H. Y. Hsu, M. T. Yeh, C. C. Chen, C. Y. Huang, Y. H. Chung, Z. F. Chang, W. C. Kuo, N. L. Chan, J. H. Weng, B. C. Chung, Y. J. Chen, C. B. Jian, C. C. Shen, H. C. Tai, S. Y. Sheu, and J. M. Fang. Chemical Inhibition of Human Thymidylate Kinase and Structural Insights into the Phosphate Binding Loop and Ligand-Induced Degradation. *Journal of Medicinal Chemistry*, 59(21):9906–9918, 2016. doi: 10.1021/acs.jmedchem.6b01280.

Q. Cui, W. S. Shin, Y. Luo, J. Tian, H. Cui, and D. Yin. Thymidylate Kinase: An Old Topic Brings New Perspectives. *Current Medicinal Chemistry*, 20(10):1286–1305, 2013. doi: 10.2174/0929867311320100006. URL `http://www.eurekaselect.com/openurl/content.php?genre=article&issn=0929-8673&volume=20&issue=10&spage=1286`.

C. Cummings, H. Murata, R. Koepsel, and A. J. Russell. Dramatically increased pH and temperature stability of chymotrypsin using dual block polymer-based protein engineering. *Biomacromolecules*, 15(3):763–771, 2014. doi: 10.1021/bm401575k. URL `http://pubs.acs.org/doi/abs/10.1021/bm401575k`.

C. S. Cummings, H. Murata, K. Matyjaszewski, and A. J. Russell. Polymer-Based Protein Engineering Enables Molecular Dissolution of Chymotrypsin in Acetonitrile. *ACS Macro Letters*, 5(4):493–497, Apr. 2016. ISSN 2161-1653, 2161-1653. doi: 10.1021/acsmacrolett.6b00137. URL `https://pubs.acs.org/doi/10.1021/acsmacrolett.6b00137`.

F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax. A multidimensional spectral processing system based on pipes. *Journal of Biomolecular NMR*, 6(3):277–293, 1995.

C. Dhalluin, A. Ross, L. A. Leuthold, S. Foser, B. Gsell, F. Müller, and H. Senn. Structural and biophysical characterization of the 40 kDa PEG-interferon-α2aand its individual positional isomers. *Bioconjugate Chemistry*, 16(3): 504–517, 2005. ISSN 1043-1802. doi: 10.1021/bc049781+.

G. Digilio, L. Barbero, C. Bracco, D. Corpillo, P. Esposito, G. Piquet, S. Traversa, and S. Aime. NMR structure of two novel polyethylene glycol conjugates of the human growth hormone-releasing factor, hGRF(1-29)-NH2. *Journal of the American Chemical Society*, 125(12):3458–3470, 2003. doi: 10.1021/ja021264j. URL `http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=12643708&retmode=ref&cmd=prlinks`.

R. Duncan and M. J. Vicent. Polymer therapeutics-prospects for 21st century: The end of the beginning. *Advanced Drug Delivery Reviews*, 65(1):60–70, Jan. 2013. ISSN 0169409X. doi: 10.1016/j.addr.2012.08.012. URL `https://linkinghub.elsevier.com/retrieve/pii/S0169409X12002566`.

A. E. Enciso, L. Fu, S. Lathwal, M. Olszewski, Z. Wang, S. R. Das, A. J. Russell, and K. Matyjaszewski. Biocatalytic "Oxygen-Fueled" Atom Transfer Radical Polymerization. *Angewandte Chemie International Edition*, 57(49): 16157–16161, Dec. 2018. ISSN 14337851. doi: 10.1002/anie.201809018. URL `http://doi.wiley.com/10.1002/anie.201809018`.

B. T. Farmer, R. A. Venters, L. D. Spicer, M. G. Wittekind, and L. Müller. A refocused and optimized HNCA: Increased sensitivity and resolution in large macromolecules. *Journal of Biomolecular NMR*, 2(2):195–202, Mar. 1992. ISSN 0925-2738, 1573-5001. doi: 10.1007/BF01875530. URL `http://link.springer.com/10.1007/BF01875530`.

P. Feldhaus, T. Fröhlich, R. S. Goody, M. Isakov, and R. H. Schirmer. Synthetic Inhibitors of Adenylate Kinases in the Assays for ATPases and Phosphokinases. *European Journal of Biochemistry*, 57(1):197–204, 1975. ISSN 0014-2956 (Print) 0014-2956 (Linking). doi: 10.1111/j.1432-1033.1975.tb02291.x.

P. a. Furman, J. a. Fyfe, M. H. St Clair, K. Weinhold, J. L. Rideout, G. a. Freeman, S. N. Lehrman, D. P. Bolognesi, S. Broder, and H. Mitsuya. Phosphorylation of 3'-azido-3'-deoxythymidine and selective interaction of the 5'-triphosphate with human immunodeficiency virus reverse transcriptase. *Proceedings of the National Academy of Sciences of the United States of America*, 83(21):8333–8337, 1986. ISSN 0027-8424 (Print) 0027-8424 (Linking). doi: 10.1073/pnas.83.21.8333. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=386922&tool=pmcentrez&rendertype=abstract`.

P. Gans, O. Hamelin, R. Sounier, I. Ayala, M. Durá, C. Amero, M. Noirclerc-Savoye, B. Franzetti, M. Plevin, and J. Boisbouvier. Stereospecific Isotopic Labeling of Methyl Groups for NMR Spectroscopic Studies of High-Molecular-Weight Proteins. *Angewandte Chemie International Edition*, 49 (11):1958–1962, Mar. 2010. ISSN 14337851. doi: 10.1002/anie.200905660. URL `http://doi.wiley.com/10.1002/anie.200905660`.

J. Garcia de la Torre, M. Huertas, and B. Carrasco. HYDRONMR: Prediction of NMR Relaxation of Globular Proteins from Atomic-Level Structures and Hydrodynamic Calculations. *Journal of Magnetic Resonance*, 147(1):138–146, Nov. 2000. ISSN 10907807. doi: 10.1006/jmre.2000.2170. URL `https://linkinghub.elsevier.com/retrieve/pii/S1090780700921700`.

M. Gauthier, A. and H.-A. Klok. Polymer-protein conjugates: an enzymatic activity perspective. *Polymer Chemistry*, 1(9):1352–1373, 2010. ISSN 1759-9954, 1759-9962. doi: 10.1039/c0py00215a. URL `http://xlink.rsc.org/?DOI=c0py00215a`.

N. K. Goto, K. H. Gardner, G. A. Mueller, R. C. Willis, and L. E. Kay. A robust and cost-effective method for the production of Val, Leu, Ile (δ1) methyl-protonated 15n-, 13c-, 2h-labeled proteins. *Journal of Biomolecular NMR*, 13(4):369–374, 1999. ISSN 09252738. doi: 10.1023/A:1008393201236. URL `http://link.springer.com/10.1023/A:1008393201236`.

G. N. Grover and H. D. Maynard. Protein–polymer conjugates: synthetic approaches by controlled radical polymerizations and interesting applications. *Current Opinion in Chemical Biology*, 14(6):818–827, Dec. 2010. ISSN 13675931. doi: 10.1016/j.cbpa.2010.10.008. URL `https://linkinghub.elsevier.com/retrieve/pii/S1367593110001493`.

S. Grzesiek and A. Bax. Improved 3d triple-resonance NMR techniques applied to a 31 kDa protein. *Journal of Magnetic Resonance (1969)*, 96(2):432–440, Feb. 1992. ISSN 00222364. doi: 10.1016/0022-2364(92)90099-S. URL `https://linkinghub.elsevier.com/retrieve/pii/002223649290099S`.

X. Han, A. Aslanian, and J. R. Yates. Mass spectrometry for proteomics. *Current Opinion in Chemical Biology*, 12(5):483–490, Oct. 2008. ISSN 13675931. doi: 10.1016/j.cbpa.2008.07.024. URL `https://linkinghub.elsevier.com/retrieve/pii/S1367593108001178`.

D. F. Hansen, P. Vallurupalli, and L. E. Kay. Using relaxation dispersion NMR spectroscopy to determine structures of excited, invisible protein states. *Journal of Biomolecular NMR*, 41(3):113–120, July 2008a. ISSN 0925-2738, 1573-5001. doi: 10.1007/s10858-008-9251-5. URL http://link.springer.com/10.1007/s10858-008-9251-5.

D. F. Hansen, P. Vallurupalli, P. Lundström, P. Neudecker, and L. E. Kay. Probing Chemical Shifts of Invisible States of Proteins with Relaxation Dispersion NMR Spectroscopy: How Well Can We Do? *Journal of the American Chemical Society*, 130(8):2667–2675, Feb. 2008b. ISSN 0002-7863, 1520-5126. doi: 10.1021/ja078337p. URL https://pubs.acs.org/doi/10.1021/ja078337p.

C.-M. Hu, M.-T. Yeh, N. Tsao, C.-W. Chen, Q.-Z. Gao, C.-Y. Chang, M.-H. Lee, J.-M. Fang, S.-Y. Sheu, C.-J. Lin, M.-C. Tseng, Y.-J. Chen, and Z.-F. Chang. Tumor Cells Require Thymidylate Kinase to Prevent dUTP Incorporation during DNA Repair. *Cancer Cell*, 22(1):36–50, July 2012. ISSN 15356108. doi: 10.1016/j.ccr.2012.04.038. URL http://linkinghub.elsevier.com/retrieve/pii/S153561081200178X.

C.-M. Hu, N. Tsao, Y.-T. Wang, Y.-J. Chen, and Z.-F. Chang. Thymidylate kinase is critical for DNA repair *via* ATM-dependent Tip60 complex formation. *The FASEB Journal*, 33(2):2017–2025, Feb. 2019. ISSN 0892-6638, 1530-6860. doi: 10.1096/fj.201800856R. URL https://www.fasebj.org/doi/10.1096/fj.201800856R.

C.-Y. Huang, Y.-C. Chen, B. A. Wu-Hsieh, J.-M. Fang, and Z.-F. Chang. The Ca-loop in thymidylate kinase is critical for growth and contributes to pyrimidine drug sensitivity of *Candida albicans*. *Journal of Biological Chemistry*, 294(27):10686–10697, July 2019. ISSN 0021-9258, 1083-351X. doi: 10.1074/jbc.RA118.006798. URL http://www.jbc.org/lookup/doi/10.1074/jbc.RA118.006798.

S.-H. Huang, A. Tang, B. Drisco, S.-Q. Zhang, R. Seeger, C. Li, and A. Jong. Human dTMP Kinase: Gene Expression and Enzymatic Activity Coinciding with Cell Cycle Progression and Cell Growth. *DNA and Cell Biology*, 13(5):461–471, May 1994. ISSN 1044-5498, 1557-7430. doi: 10.1089/dna.1994.13.461. URL http://www.liebertpub.com/doi/10.1089/dna.1994.13.461.

W. Jakubowski and K. Matyjaszewski. Activator Generated by Electron Transfer for Atom Transfer Radical Polymerization. *Macromolecules*, 38(10):4139–4146, May 2005. ISSN 0024-9297, 1520-5835. doi: 10.1021/ma047389l. URL https://pubs.acs.org/doi/10.1021/ma047389l.

A. Jordan and P. Reichard. Ribonucleotide Reductases. *Annual Review of Biochemistry*, 67(1):71–98, June 1998. ISSN 0066-4154, 1545-4509. doi: 10.1146/annurev.biochem.67.1.71. URL http://www.annualreviews.org/doi/10.1146/annurev.biochem.67.1.71.

A. Kamah, I. Huvent, F.-X. Cantrelle, H. Qi, G. Lippens, I. Landrieu, and C. Smet-Nocca. Nuclear Magnetic Resonance Analysis of the Acetylation Pattern of the Neuronal Tau Protein. *Biochemistry*, 53(18):3020–3032, 2014. doi: 10.1021/bi500006v. URL `http://pubs.acs.org/doi/abs/10.1021/bi500006v`.

B. Kaupbayeva, H. Murata, A. Lucas, K. Matyjaszewski, J. S. Minden, and A. J. Russell. Molecular Sieving on the Surface of a Nano-Armored Protein. *Biomacromolecules*, 20(3):1235–1245, Mar. 2019. ISSN 1525-7797, 1526-4602. doi: 10.1021/acs.biomac.8b01651. URL `https://pubs.acs.org/doi/10.1021/acs.biomac.8b01651`.

L. E. Kay, M. Ikura, R. Tschudin, and A. Bax. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *Journal of Magnetic Resonance (1969)*, 89(3):496–514, Oct. 1990. ISSN 00222364. doi: 10.1016/0022-2364(90)90333-5. URL `https://linkinghub.elsevier.com/retrieve/pii/0022236490903335`.

P.-Y. Ke. Control of dTTP pool size by anaphase promoting complex/cyclosome is essential for the maintenance of genetic stability. *Genes & Development*, 19(16):1920–1933, Aug. 2005. ISSN 0890-9369. doi: 10.1101/gad.1322905. URL `http://www.genesdev.org/cgi/doi/10.1101/gad.1322905`.

S. J. Kerns, R. V. Agafonov, Y. J. Cho, F. Pontiggia, R. Otten, D. V. Pachov, S. Kutter, L. A. Phung, P. N. Murphy, V. Thai, T. Alber, M. F. Hagan, and D. Kern. The energy landscape of adenylate kinase during catalysis. *Nature Structural and Molecular Biology*, 22(2):124–131, 2015. ISSN 0044122374014. doi: 10.1038/nsmb.2941. URL `http://dx.doi.org/10.1038/nsmb.2941`.

I. R. Kleckner and M. P. Foster. An introduction to NMR-based approaches for measuring protein dynamics. *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1814(8):942–968, 2011. ISSN 1570-9639. doi: 10.1016/j.bbapap.2010.10.012. URL `http://dx.doi.org/10.1016/j.bbapap.2010.10.012`.

I. R. Kleckner and M. P. Foster. GUARDD: user-friendly MATLAB software for rigorous analysis of CPMG RD NMR data. *Journal of Biomolecular NMR*, 52(1):11–22, Jan. 2012. ISSN 0925-2738, 1573-5001. doi: 10.1007/s10858-011-9589-y. URL `http://link.springer.com/10.1007/s10858-011-9589-y`.

E. R. Kline, L. Bassit, B. I. Hernandez-Santiago, M. A. Detorio, B. Liang, D. J. Kleinhenz, E. R. Walp, S. Dikalov, D. P. Jones, R. F. Schinazi, and R. L. Sutliff. Long-Term Exposure to AZT, but not d4t, Increases Endothelial Cell Oxidative Stress and Mitochondrial Dysfunction. *Cardiovascular Toxicology*, 9(1):1–12, Mar. 2009. ISSN 1530-7905, 1559-0259. doi: 10.1007/s12012-008-9029-8. URL `http://link.springer.com/10.1007/s12012-008-9029-8`.

H. Koss, M. Rance, and A. G. Palmer. General Expressions for Carr–Purcell–Meiboom–Gill Relaxation Dispersion for N -Site Chemical Exchange. *Biochemistry*, 57(31):4753–4763, Aug. 2018. ISSN 0006-2960, 1520-4995. doi: 10.1021/acs.biochem.8b00370. URL `http://pubs.acs.org/doi/10.1021/acs.biochem.8b00370`.

E. L. Kovrigin, J. G. Kempf, M. J. Grey, and J. P. Loria. Faithful estimation of dynamics parameters from CPMG relaxation dispersion measurements. *Journal of Magnetic Resonance*, 180(1):93–104, May 2006. ISSN 10907807. doi: 10.1016/j.jmr.2006.01.010. URL `https://linkinghub.elsevier.com/retrieve/pii/S1090780706000164`.

G. S. Kumar, M. W. Clarkson, M. B. A. Kunze, D. Granata, A. J. Wand, K. Lindorff-Larsen, R. Page, and W. Peti. Dynamic activation and regulation of the mitogen-activated protein kinase p38. *Proceedings of the National Academy of Sciences*, page 201721441, Apr. 2018. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1721441115. URL `http://www.pnas.org/lookup/doi/10.1073/pnas.1721441115`.

A. Lavie, I. Schlichting, I. R. Vetter, M. Konrads, J. Reinstein, and R. S. Goody. The bottleneck in AZT activation. *Nature Medicine*, 3(8):922–924, Aug. 1997. ISSN 1078-8956, 1546-170X. doi: 10.1038/nm0897-922. URL `http://www.nature.com/articles/nm0897-922`.

A. Lavie, M. Konrad, R. Brundiers, R. S. Goody, I. Schlichting, and J. Reinstein. Crystal Structure of Yeast Thymidylate Kinase Complexed with the Bisubstrate Inhibitor P^1 -(5‘-Adenosyl) P^5 -(5‘-Thymidyl) Pentaphosphate (TP $_5$ A) at 2.0 Å Resolution: Implications for Catalysis and AZT Activation † , ‡. *Biochemistry*, 37(11):3677–3686, Mar. 1998. ISSN 0006-2960, 1520-4995. doi: 10.1021/bi9720787. URL `http://pubs.acs.org/doi/abs/10.1021/bi9720787`.

J. H. Lee, J. Ying, and A. Bax. Nuclear Magnetic Resonance Observation of α-Synuclein Membrane Interaction by Monitoring the Acetylation Reactivity of Its Lysine Side Chains. *Biochemistry*, 55(35):4949–4959, 2016. doi: 10.1021/acs.biochem.6b00637. URL `http://pubs.acs.org/doi/abs/10.1021/acs.biochem.6b00637`.

I. Li de la Sierra, H. Munier-Lehmann, A. Gilles, O. Bârzu, and M. Delarue. X-ray structure of TMP kinase from Mycobacterium tuberculosis complexed with TMP at 1.95 Å resolution. *Journal of Molecular Biology*, 311(1):87–100, Aug. 2001. ISSN 00222836. doi: 10.1006/jmbi.2001.4843. URL `http://linkinghub.elsevier.com/retrieve/pii/S0022283601948430`.

G. P. Lisi and J. P. Loria. Solution NMR Spectroscopy for the Study of Enzyme Allostery. *Chemical Reviews*, 116(11):6323–6369, June 2016. ISSN 0009-2665, 1520-6890. doi: 10.1021/acs.chemrev.5b00541. URL `http://pubs.acs.org/doi/10.1021/acs.chemrev.5b00541`.

Y. Liu, K. Marks, G. S. Cowley, J. Carretero, Q. Liu, T. J. F. Nieland, C. Xu, T. J. Cohoon, P. Gao, Y. Zhang, Z. Chen, A. B. Altabef, J. H. Tchaicha, X. Wang, S. Choe, E. M. Driggers, J. Zhang, S. T. Bailey, N. E. Sharpless, D. N. Hayes, N. M. Patel, P. A. Janne, N. Bardeesy, J. A. Engelman, B. D. Manning, R. J. Shaw, J. M. Asara, R. Scully, A. Kimmelman, L. A. Byers, D. L. Gibbons, I. I. Wistuba, J. V. Heymach, D. J. Kwiatkowski, W. Y. Kim, A. L. Kung, N. S. Gray, D. E. Root, L. C. Cantley, and K.-K. Wong. Metabolic and Functional Genomic Studies Identify Deoxythymidylate Kinase as a Target in LKB1-Mutant Lung Cancer. *Cancer Discovery*, 3(8):870–879, Aug. 2013. ISSN 2159-8274, 2159-8290. doi: 10.1158/2159-8290.CD-13-0015. URL `http://cancerdiscovery.aacrjournals.org/cgi/doi/10.1158/2159-8290.CD-13-0015`.

A. Lucas, B. Kaupbayeva, H. Murata, C. Cummings, A. J. Russell, and J. S. Minden. Utilization of the Polymer Sieving Effect for the Removal of the Small Molecule Biotin-CDM. *ACS Applied Polymer Materials*, 1(11):2897–2906, Nov. 2019. ISSN 2637-6105, 2637-6105. doi: 10.1021/acsapm.9b00611. URL `https://pubs.acs.org/doi/10.1021/acsapm.9b00611`.

M. Lucius, R. Falatach, C. McGlone, and K. Makaroff. Investigating the Impact of Polymer Functional Groups on the Stability and Activity of Lysozyme–Polymer Conjugates. *Biomacromolecules*, 17(3):1123–1134, 2016. doi: 10.1021/acs.biomac.5b01743. URL `http://pubs.acs.org/doi/abs/10.1021/acs.biomac.5b01743`.

Z. Luz and S. Meiboom. Nuclear Magnetic Resonance Study of the Protolysis of Trimethylammonium Ion in Aqueous Solution—Order of the Reaction with Respect to Solvent. *The Journal of Chemical Physics*, 39(2):366–370, July 1963. ISSN 0021-9606, 1089-7690. doi: 10.1063/1.1734254. URL `http://aip.scitation.org/doi/10.1063/1.1734254`.

R. J. Mancini, J. Lee, and H. D. Maynard. Trehalose Glycopolymers for Stabilization of Protein Conjugates to Environmental Stressors. *Journal of the American Chemical Society*, 134(20):8474–8479, May 2012. ISSN 0002-7863, 1520-5126. doi: 10.1021/ja2120234. URL `https://pubs.acs.org/doi/10.1021/ja2120234`.

O. Millet, J. P. Loria, C. D. Kroenke, M. Pons, and A. G. Palmer. The Static Magnetic Field Dependence of Chemical Exchange Linebroadening Defines the NMR Chemical Shift Time Scale. *Journal of the American Chemical Society*, 122(12):2867–2877, Mar. 2000. ISSN 0002-7863, 1520-5126. doi: 10.1021/ja993511y. URL `http://pubs.acs.org/doi/abs/10.1021/ja993511y`.

D. Moatsou, J. Li, A. Ranji, A. Pitto-Barry, I. Ntai, M. C. Jewett, and R. K. O'Reilly. Self-Assembly of Temperature-Responsive Protein–Polymer Bioconjugates. *Bioconjugate chemistry*, 26(9):1890–1899, 2015. doi: 10.1021/acs.bioconjchem.5b00264. URL `http://pubs.acs.org/doi/10.1021/acs.bioconjchem.5b00264`.

H. Murata, C. S. Cummings, R. R. Koepsel, and A. J. Russell. Polymer-Based Protein Engineering Can Rationally Tune Enzyme Activity, pH-Dependence, and Stability. *Biomacromolecules*, 14(6):1919–1926, June 2013. ISSN 1525-7797, 1526-4602. doi: 10.1021/bm4002816. URL `https://pubs.acs.org/doi/10.1021/bm4002816`.

N. Ostermann, A. Lavie, S. Padiyar, R. Brundiers, T. Veit, J. Reinstein, R. S. Goody, M. Konrad, and I. Schlichting. Potentiating AZT activation: structures of wild-type and mutant human thymidylate kinase suggest reasons for the mutants' improved kinetics with the HIV prodrug metabolite AZTMP. *Journal of Molecular Biology*, 304(1):43–53, 2000a. doi: 10.1006/jmbi.2000.4175. URL `http://linkinghub.elsevier.com/retrieve/pii/S0022283600941755`.

N. Ostermann, I. Schlichting, R. Brundiers, M. Konrad, J. Reinstein, T. Veit, R. S. Goody, and A. Lavie. Insights into the phosphoryltransfer mechanism of human thymidylate kinase gained from crystal structures of enzyme complexes along the reaction coordinate. *Structure*, 8(6):629–642, 2000b. doi: 10.1016/s0969-2126(00)00149-0. URL `http://linkinghub.elsevier.com/retrieve/pii/S0969212600001490`.

N. Ostermann, D. Segura-Peña, C. Meier, T. Veit, C. Monnerjahn, M. Konrad, and A. Lavie. Structures of Human Thymidylate Kinase in Complex with Prodrugs: Implications for the Structure-Based Design of Novel Compounds †. *Biochemistry*, 42(9):2568–2577, 2003. doi: 10.1021/bi027302t. URL `http://pubs.acs.org/doi/abs/10.1021/bi027302t`.

A. G. Palmer and H. Koss. Chemical Exchange. In *Methods in Enzymology*, volume 615, pages 177–236. Elsevier, 2019. ISBN 978-0-12-816762-5. doi: 10.1016/bs.mie.2018.09.028. URL `https://linkinghub.elsevier.com/retrieve/pii/S0076687918303914`.

S. J. Perkins and K. Wüthrich. Ring current effects in the conformation dependent NMR chemical shifts of aliphatic protons in the basic pancreatic trypsin inhibitor. *Biochimica et Biophysica Acta (BBA) - Protein Structure*, 576(2):409–423, Feb. 1979. ISSN 00052795. doi: 10.1016/0005-2795(79)90416-1. URL `https://linkinghub.elsevier.com/retrieve/pii/0005279579904161`.

J. J. Pitt. Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *The Clinical Biochemist. Reviews*, 30(1):19–34, Feb. 2009. ISSN 0159-8090.

Y. Qi and A. Chilkoti. Protein–polymer conjugation — moving beyond PEGylation. *Current Opinion in Chemical Biology*, 28:181–193, Oct. 2015. ISSN 13675931. doi: 10.1016/j.cbpa.2015.08.009. URL `https://linkinghub.elsevier.com/retrieve/pii/S1367593115001003`.

P. Reichard. Interactions Between Deoxyribonucleotide and DNA Synthesis. *Annual Review of Biochemistry*, 57:349–374, 1988.

T. Sato, A. Neschadim, A. Lavie, T. Yanagisawa, and J. A. Medin. The Engineered Thymidylate Kinase (TMPK)/AZT Enzyme-Prodrug Axis Offers Efficient Bystander Cell Killing for Suicide Gene Therapy of Cancer. *PLoS ONE*, 8(10):e78711, Oct. 2013. ISSN 1932-6203. doi: 10.1371/journal.pone.0078711. URL `https://dx.plos.org/10.1371/journal.pone.0078711`.

K. Scheffzek, W. Kliche, L. Wiesmüller, and J. Reinstein. Crystal Structure of the Complex of UMP/CMP Kinase from *Dictyostelium discoideum* and the Bisubstrate Inhibitor P^{1} -(5‘-Adenosyl) P^{5} -(5‘-Uridyl) Pentaphosphate ($UP_{5}A$) and Mg^{2+} at 2.2 Å: Implications for Water-Mediated Specificity ‡. *Biochemistry*, 35(30):9716–9727, Jan. 1996. ISSN 0006-2960, 1520-4995. doi: 10.1021/bi960642s. URL `https://pubs.acs.org/doi/10.1021/bi960642s`.

H. Schellekens, W. E. Hennink, and V. Brinks. The Immunogenicity of Polyethylene Glycol: Facts and Fiction. *Pharmaceutical Research*, 30(7):1729–1734, July 2013. ISSN 0724-8741, 1573-904X. doi: 10.1007/s11095-013-1067-7. URL `http://link.springer.com/10.1007/s11095-013-1067-7`.

K. Sinha and G. S. Rule. The Structure of Thymidylate Kinase from *Candida albicans* Reveals a Unique Structural Element. *Biochemistry*, 56(33):4360–4370, Aug. 2017. ISSN 0006-2960, 1520-4995. doi: 10.1021/acs.biochem.7b00498. URL `http://pubs.acs.org/doi/10.1021/acs.biochem.7b00498`.

C. Smet-Nocca, J.-M. Wieruszeski, O. Melnyk, and A. Benecke. NMR-based detection of acetylation sites in peptides. *Journal of peptide science : an official publication of the European Peptide Society*, 16(8):414–423, 2010. doi: 10.1002/psc.1257. URL `http://doi.wiley.com/10.1002/psc.1257`.

F. W. Studier. Protein production by auto-induction in high density shaking cultures. *Protein expression and purification*, 41(1):207–234, 2005. URL `http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=15915565&retmode=ref&cmd=prlinks`.

K. Sugase, T. Konuma, J. C. Lansing, and P. E. Wright. Fast and accurate fitting of relaxation dispersion data using the flexible software package GLOVE. *Journal of Biomolecular NMR*, 56(3):275–283, 2013. ISSN doi:10.1007/s10858-013-9747-5. doi: 10.1007/s10858-013-9747-5.

F. X. Theillet, C. Smet-Nocca, S. Liokatis, R. Thongwichian, J. Kosten, M. K. Yoon, R. W. Kriwacki, I. Landrieu, G. Lippens, and P. Selenko. Cell signaling, post-translational protein modifications and NMR spectroscopy. *Journal of Biomolecular NMR*, 54(3):217–236, 2012. ISSN 1573-5001 (Electronic)\r0925-2738 (Linking). doi: 10.1007/s10858-012-9674-x.

V. Tugarinov and L. E. Kay. Ile, Leu, and Val Methyl Assignments of the 723-Residue Malate Synthase G Using a New Labeling Strategy and Novel NMR Methods. *Journal of the American Chemical Society*, 125(45):13868–13878, Nov. 2003. ISSN 0002-7863, 1520-5126. doi: 10.1021/ja030345s. URL `https://pubs.acs.org/doi/10.1021/ja030345s`.

V. Tugarinov and L. E. Kay. Quantitative ^{13} C and ^{2} H NMR Relaxation Studies of the 723-Residue Enzyme Malate Synthase G Reveal a Dynamic Binding Interface †. *Biochemistry*, 44(49):15970–15977, Dec. 2005. ISSN 0006-2960, 1520-4995. doi: 10.1021/bi0519809. URL `https://pubs.acs.org/doi/10.1021/bi0519809`.

S. van der Walt, S. C. Colbert, and G. Varoquaux. The NumPy Array: A Structure for Efficient Numerical Computation. *Computing in Science & Engineering*, 13(2):22–30, Mar. 2011. ISSN 1521-9615. doi: 10.1109/MCSE.2011.37. URL `http://ieeexplore.ieee.org/document/5725236/`.

S. van Dongen, M. Nallani, J. J. L. M. Cornelissen, R. Nolte, and J. van Hest. A Three-Enzyme Cascade Reaction through Positional Assembly of Enzymes in a Polymersome Nanoreactor. *Chemistry - A European Journal*, 15(5):1107–1114, Jan. 2009. ISSN 09476539. doi: 10.1002/chem.200802114. URL `http://doi.wiley.com/10.1002/chem.200802114`.

P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, I. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and S. . . Contributors. SciPy 1.0–Fundamental Algorithms for Scientific Computing in Python. *arXiv:1907.10121 [physics]*, July 2019. URL `http://arxiv.org/abs/1907.10121`. arXiv: 1907.10121.

G. Vuister, G. Clore, A. Gronenborn, R. Powers, D. Garrett, R. Tschudin, and A. Bax. Increased Resolution and Improved Spectral Quality in Four-Dimensional 13c/13c-Separated HMQC-NOESY-HMQC Spectra Using Pulsed Field Gradients. *Journal of Magnetic Resonance, Series B*, 101(2):210–213, Apr. 1993. ISSN 10641866. doi: 10.1006/jmrb.1993.1035. URL `https://linkinghub.elsevier.com/retrieve/pii/S1064186683710356`.

Y. Wada and M. Kadoya. In-gel digestion with endoproteinase LysC. *Journal of Mass Spectrometry*, 38(1):117–118, 2003. ISSN 1076-5174 (Print). doi: 10.1002/jms.384.

J.-S. Wang and K. Matyjaszewski. Controlled/"living" radical polymerization. atom transfer radical polymerization in the presence of transition-metal complexes. *Journal of the American Chemical Society*, 117(20):5614–5615, May 1995. ISSN 0002-7863. doi: 10.1021/ja00125a035. URL `https://pubs.acs.org/doi/abs/10.1021/ja00125a035`.

Y. S. Wang, S. Youngster, J. Bausch, R. Zhang, C. McNemar, and D. F. Wyss. Identification of the major positional isomer of pegylated interferon alpha-2b. *Biochemistry*, 39(35):10634–10640, 2000. ISSN 0006-2960 (Print)\n0006-2960 (Linking). doi: 10.1021/bi000617t.

C. A. Waudby, A. Ramos, L. D. Cabrita, and J. Christodoulou. Two-Dimensional NMR Lineshape Analysis. *Scientific Reports*, 6(1), July 2016. ISSN 2045-2322. doi: 10.1038/srep24826. URL `http://www.nature.com/articles/srep24826`.

J. Whittingham, J. Carrero-Lerida, J. Brannigan, L. Ruiz-Perez, A. Silva, M. Fogg, A. Wilkinson, I. Gilbert, K. Wilson, and D. González-Pacanowska. Structural basis for the efficient phosphorylation of AZT-MP (3'-azido-3'-deoxythymidine monophosphate) and dGMP by Plasmodium falciparum type I thymidylate kinase. *Biochemical Journal*, 428(3):499–509, 2010. ISSN 1470-8728 (Electronic)\r0264-6021 (Linking). doi: 10.1042/BJ20091880. URL `http://biochemj.org/lookup/doi/10.1042/BJ20091880`.

M. P. Williamson. Using chemical shift perturbation to characterise ligand binding. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 73:1–16, 2013. ISSN 0079-6565. doi: 10.1016/j.pnmrs.2013.02.001. URL `http://dx.doi.org/10.1016/j.pnmrs.2013.02.001`.

B. M. Wohrl. Expressing engineered thymidylate kinase variants in human cells to improve AZT phosphorylation and human immunodeficiency virus inhibition. *Journal of General Virology*, 86(3):757–764, Mar. 2005. ISSN 0022-1317, 1465-2099. doi: 10.1099/vir.0.80529-0. URL `https://www.microbiologyresearch.org/content/journal/jgv/10.1099/vir.0.80529-0`.

M. Wolf-Watz, V. Thai, K. Henzler-Wildman, G. Hadjipavlou, E. Z. Eisenmesser, and D. Kern. Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nature Structural & Molecular Biology*, 11(10):945–949, Oct. 2004. ISSN 1545-9993, 1545-9985. doi: 10.1038/nsmb821. URL `http://www.nature.com/articles/nsmb821`.

J.-P. Yan, D. D. Ilsley, C. Frohlick, R. Steet, E. T. Hall, R. D. Kuchta, and P. Melan¸on. 3'-Azidothymidine (Zidovudine) Inhibits Glycosylation and Dramatically Alters Glycosphingolipid Synthesis in Whole Cells at Clinically Relevant Concentrations. *Journal of Biological Chemistry*, 270(39):22836–22841, Sept. 1995. ISSN 0021-9258, 1083-351X. doi: 10.1074/jbc.270.39.22836. URL `http://www.jbc.org/lookup/doi/10.1074/jbc.270.39.22836`.

J. Ådén, A. Verma, A. Schug, and M. Wolf-Watz. Modulation of a pre-existing conformational equilibrium tunes adenylate kinase activity. *Journal of the American Chemical Society*, 134(40):16562–16570, 2012. ISSN 1520-5126 (Electronic)\n0002-7863 (Linking). doi: 10.1021/ja3032482.

J. Ådén, C. F. Weise, K. Brännström, A. Olofsson, and M. Wolf-Watz. Structural topology and activation of an initial adenylate kinase-substrate complex. *Biochemistry*, 52(6):1055–1061, 2013. ISSN 1520-4995 (Electronic)\n0006-2960 (Linking). doi: 10.1021/bi301460k.