Understanding the Binding Mechanism of an SH3 Domain Using NMR and ITC

by

Jean-Philippe Demers

Department of Chemistry
McGill University, Montreal

October 2008

A thesis submitted to McGill University
in partial fulfilment of the requirements of the degree of
Master of Science

ABSTRACT

The transient interaction of the Fyn SH3 domain with a proline-rich peptide was studied using Nuclear Magnetic Resonance (NMR) and isothermal titration calorimetry (ITC). Our results are consistent with an effectively two-state binding reaction mechanism. The association rate constants calculated using NMR and ITC data exhibit the magnitude and temperature dependence consistent with a diffusion-limited process. An Eyring plot of k_{off} exhibits a slight curvature, suggesting that non-polar residues of the hydrophobic interface are solvated in the transition state for dissociation. The temperature dependence of NMR-derived dissociation rates closely matches the heat capacity and enthalpy changes determined by ITC. The combination of NMR and ITC data sheds light on how the Fyn tyrosine kinase is activated by binding to proline-rich targets, and represents a powerful approach for characterizing transient protein-protein interactions.

KEYWORDS. Nuclear Magnetic Resonance, Isothermal Titration Calorimetry, SH3 Domain, Binding Kinetics

RÉSUMÉ

L'interaction rapide entre le domaine SH3 de la protéine Fyn et un peptide riche en proline a été étudiée par résonance magnétique nucléaire (RMN) et titration calorimétrique isotherme (TCI). Nos résultats concordent avec un mécanisme de réaction à deux états.

Les vitesses d'association, calculées par RMN et TCI, indiquent une réaction limitée par la diffusion. Dans un graphique d'Eyring, les vitesses de dissociation exhibent une légère courbure, suggérant que l'interface hydrophobique entre les deux composés est solvatés dans l'état de transition; cette courbure est apparentée à la différence de capacité thermique ayant été déterminée par TCI. L'utilisation combinée de RMN et TCI nous a permis de mieux comprendre comment la tyrosine kinase Fyn est activée par son association avec des peptides. Ainsi, nous avons développé une approche robuste pour caractériser les interactions rapides entre protéines.

MOTS-CLÉS: Résonance magnétique nucléaire, titration calorimétrique isotherme, domaine SH3, cinétique d'interaction

ACKNOWLEDGEMENTS

I would like to thank the people who gave me invaluable help for the realisation of this project. Lee Freiburger trained me on how to carry out ITC experiments. Dr Tara Sprules taught me how to record NMR experiments. Imran Sulemankhil, Olivia Dinica, Karen Galoustov and Erick Meneses assisted me in various laboratory tasks, notably the purification of the Fyn SH3 domain. Dr Lewis Kay provided insightful discussions on the project, as well as a new NMR pulse sequence. When in critical need of laboratory help, the doors of the groups of Dr Kalle Gehring and Dr Karine Auclair have always been open. Dr Alan Davidson provided the plasmid used in those studies.

A special gratitude is dedicated to my research team for the fruitful discussions and their constant support in many aspects of what constitutes the graduate life: Patrick Farber, Teresa Miletti, Hariyanto Darmawan, Karen Galoustov, Erick Meneses and Deniz Aydin.

For their support, I want to thank my family: Diane, Roland Pierre and Karoline; and friends, particularly Zhijie Chua, Zhuo Li, Jun Wang and Dr Hsiu-Fang Fan who shared my path when I entered graduate school in Chemistry. Zhijie Chua pushed his dedication to the level of proof-reading this manuscript.

Foremost, I want to thank my supervisor, Dr Anthony Mittermaier, a man of excellence and dedication in science, and exceptional human qualities. Dr Mittermaier has been caring and motivating, such that this project was a nurturing and exciting experience.

I want to dedicate this work to past supervisors and mentors, who lead me to pursue my interests in science, and which I hope to emulate.

TABLE OF CONTENT

The present work has been submitted for publication in *J. Am. Chem. Soc.*. I, Jean-Philippe Demers, did all the experimental work and data analysis. Dr Anthony Mittermaier assisted me in the redaction of the manuscript.

List of Figures

List of Tables

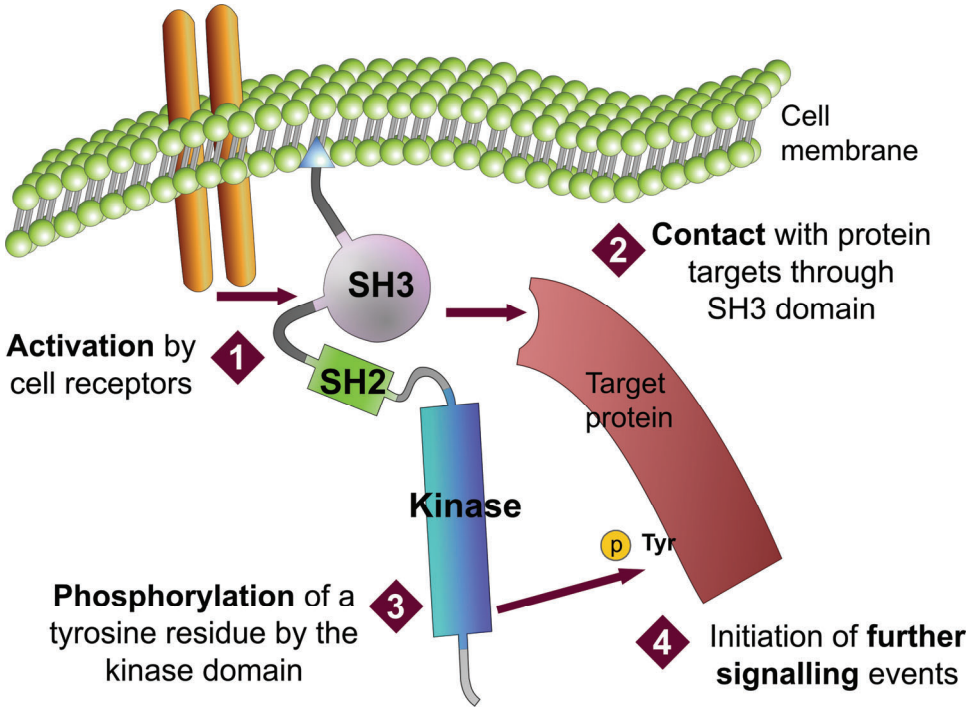
CHAPTER 1: INTRODUCTION TO PROTEIN BINDING KINETICS

1.1 Research motivation

The formation of multi-protein complexes plays a crucial role in the maintenance of homeostasis. Many cellular processes are controlled by non-covalent macromolecular interactions (Fig. 1.1).

Efforts have been made to understand protein interaction networks in terms of their equilibrium affinities[2]. However, compared to equilibrium affinity constants, rates of association and dissociation provide a more complete depiction of biological systems since *i)* organisms do not operate at equilibrium, and *ii) in vivo* processes often require the binding to be rapid rather than tight[3].

Figure 1.1 Role of Fyn in cell signalling

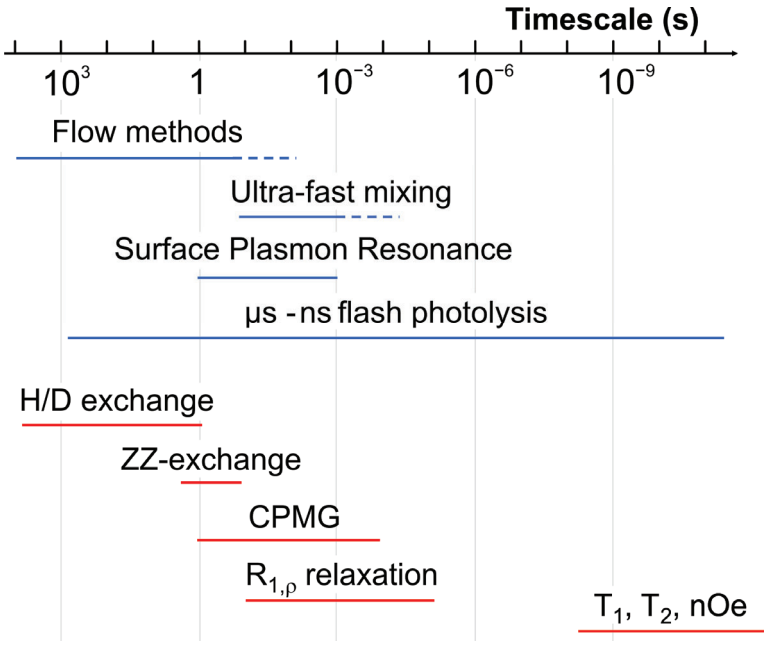


The quantification of kinetic and thermodynamic interaction parameters can bring about a keener understanding of the regulation of biological processes. This requires knowledge of the factors that influence the rates of macromolecular association and dissociation. Accordingly, the objective of the current project is to characterize the mechanism of a SH3 domain-peptide interaction by Nuclear Magnetic Resonance (NMR) and Isothermal Titration Calorimetry (ITC). Our model system is novel in two regards: it involves fast association events, and makes use of an unstructured peptide ligand.

Most highly-studied systems bind very tightly and dissociate very slowly, with bound lifetimes on the order of hours or days[4]. Much less is known about weaker complexes with bound lifetimes on the order of milliseconds. Association and dissociation events happening on those fast timescales cannot be resolved with conventional techniques (Fig. 1.2). We employed NMR methods that have been recently developed to study protein excited states[4-6], which are higher in free energy than the ground state, and applied those methodologies to the problem of protein-peptide binding.

A short peptide was selected for our studies (see Table 1.1). Small peptides are likely to be intrinsically unstructured and to have a larger conformational freedom in their free than in their bound state. When these molecules enter the bound state, they have to rapidly select the correct conformations and relative orientations for productive interaction from the landscape of different options. Numerous theoretical, computational and experimental studies have addressed how this process takes place[2,7-11], however there are still some aspects of the process that are not well understood.

Figure 1.2 Techniques to study fast-timescale reactions



Time resolution for several kinetics measurement techniques. Blue lines correspond to conventional techniques and red lines correspond to NMR-based techniques. Adapted from Laidler *et al*, 2003[5].

1.2 Why use NMR?

NMR presents several advantages over other techniques, and is well-suited for the aim of this project. At first, there is no dead-time for NMR measurements of binding rates, since the experiment is conducted under dynamic equilibrium conditions. In contrast, ultra-fast mixing techniques, either continuous flow or stopped flow, usually have a dead time in the low micro-second range[6] (Fig. 1.2), in which time rapid association events can already take place.

Recently-developed NMR experiments, Carr-Purcell-Meiboom-Gill (CPMG) and R1ρ[7], can measure exchange rates up to 10^4 s^{-1}. Kinetic information is derived from a quantitative analysis of the magnetization transfer and spectral broadening produced by the exchange between bound and free states in a partially ligand-saturated sample.

Contrary to conventional methods, in which the reaction is monitored by a single signal, $^{1}H-^{15}N$ HSQC spectra have contributions from all amides from the protein. Each resonance constitutes an independent probe of the process and provides site-specific information. Fitting the data as a group can help identify intermediate states and concerted transitions.

Finally, NMR experiments are carried out directly in the solution phase and require no chemical modifications, such as the conjugation of a fluorescent probe.

1.3 Our model system

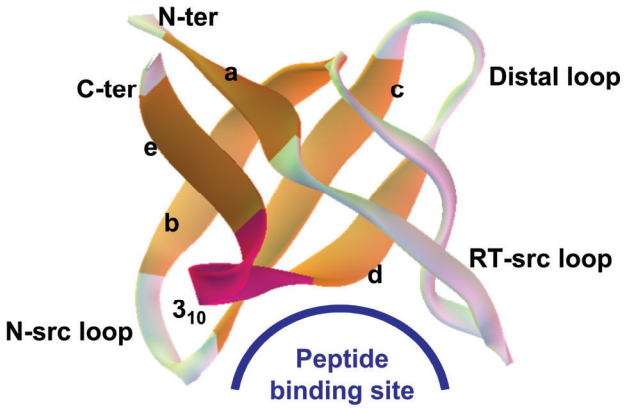
The SH3 domain fold

The SH3 domain from the Fyn tyrosine-protein kinase was selected as an experimental model system. The construct used comprises residues 81 to 148 of the chicken (*Gallus gallus*) isoform of the Fyn tyrosine kinase fused to an N-terminal 6-histidine affinity purification tag. Except for the affinity tag, the construct differs by one amino acid (E88V) from the human Fyn SH3 domain. The Fyn SH3 domain was selected for several reasons. SH3 domains are small in size, only 60 amino acids, which provides them with very favourable spectroscopic properties for NMR analysis. The Fyn SH3 domain has high stability, as demonstrated by a sharp melting transition at 70.5ºC[14], that makes it amenable for simplified purification procedures[8].

SH3 domains rank among the best studied protein folds, owing to a large number of proteins which possess this docking module. SH3 domains have been identified in over 350 proteins[9]. The SH3 fold is composed of 5 β-strands, a 3_{10}-helix and 3 extended loops (Fig. 1.3). The function of SH3 domains is to mediate the formation of multi-protein

complexes, by recognition and binding to proline-rich sequences, so as to amplify and transmit intracellular chemical signals[10] (Fig. 1.1).

Figure 1.3 Structure of the SH3 fold

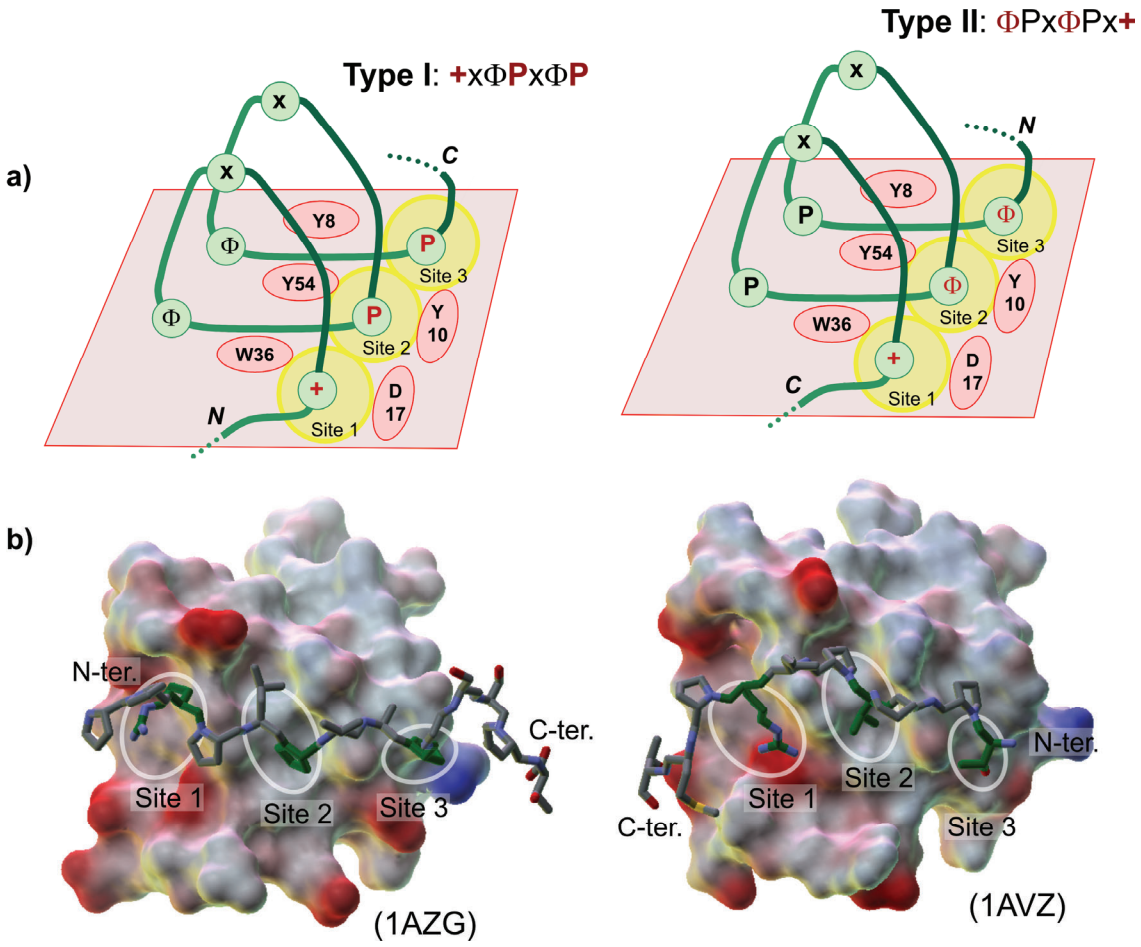


The SH3 fold is composed of 5 β-strands (yellow) forming 2 sheets, a 3_{10}-helix (magenta) and 3 extended flexible loops. The N-src loop, RT-src and 3_{10} helix form a cleft which constitutes the binding site for proline-rich ligands.

SH3 ligands and binding

SH3 domains interact promiscuously with different proline-rich sequences, which may be separated into two classes (see Fig. 1.4), based on the relative orientation of the peptide chain within the binding site: Type I, with consensus sequence +xΦPxΦP, and Type II, with sequence ΦPxΦPx+, where (+) is a positively charged residue, usually arginine, (P) is proline, and (Φ) is a hydrophobic residue[11]. The structure of the Fyn SH3 domain has been solved in its *apo* form[12], and when bound to both Type I[13] and Type II[21] ligands. The Fyn SH3 domain binds promiscuously to both classes with an affinity in the mM to μM range[20-22].

Figure 1.4 Binding modes of SH3 ligands



(a) Schematics of peptide ligands (green) bound to the surface of the SH3 domain (red). The three binding pockets (site 1, 2, 3) are identified as well as their lining residues. Type 1 ligands (+xΦ**P**xΦ**P**) bind with the N-terminus closer to site 1, while Type 2 ligands (**Φ**Px**Φ**Px+) have their C-terminus near site 1. Both types adopt a poly-proline helix type II conformation upon binding. Adapted from Dalgarno *et al*, 1997[10]. **(b)** (Left) Average NMR structure of the Fyn SH3 domain bound to a Type 1 ligand from the p85 subunit of PI3-kinase[14]. (Right) X-ray structure of the Fyn SH3 domain bound to a Type 2 ligand from HIV-1 Nef[15]. Peptide residues interacting with one of the binding pockets are highlighted in green. Proteins are represented as a contact surface, with positively charged regions in blue, neutral regions in grey, and negatively charged regions in red.

The interaction interface is composed of three pockets. The first site hosts the positive ligand side-chain through an acidic patch containing conserved Trp and Asp residues. The two other adjacent sites are predominantly non-polar (See Fig. 1.4 and 3.3) and accept the proline core of the ligand through three conserved Tyr sidechains[10]. SH3 ligands adopt a polyproline helix type II (PPII) conformation upon binding, although they are unstructured in their free form.

A determining factor governing the ligand affinity and class specificity was found to be the extent to which the motions of a conserved tryptophane residue (W36) are restricted by a neighbouring residue[16]. SH3 domain ligands and structure of their complexes have been reviewed in (Dalgarno *et al*, 1998)[10]. Some high affinity ligands for the Fyn SH3 domain have been identified in studies employing phage display[17] and combinatorial library strategies[18].

A list of Fyn SH3 domain ligands is presented in Table 1.1. For this study, we made the conservative decision of selecting the ligand with the highest affinity among this list, for peptides containing only natural amino acids. The peptide used in this study (Ac-VSLAR**R**PL**P**PL**P**-NH_2) follows the Type I consensus sequence, +xx**P**xx**P**, and was previously found to have an affinity of 0.43 μM[17]. The peptide was synthesized with N-terminal acetylation and C-terminal amidation to prevent premature degradation and to ensure a backbone polarity similar to what is found in longer peptide chains.

1.4 Measurement of reaction rates

A two-state bimolecular binding reaction consists of the association of free protein P and ligand X to form the bound complex PX, and of the dissociation of the complex to its free constituents (Eq. 1.1).

$$P \underset{k_{off}}{\overset{k_{on}[X]}{\rightleftharpoons}} PX \qquad \text{(Eq. 1.1)}$$

k_{off} is the first-order dissociation rate constant; k_{on} is the second-order association rate constant, and [X] is the concentration of uncomplexed peptide in solution.

The binding process causes changes in the NMR spectrum that can provide information about the different local chemical environments which nuclei experience, and about the process of exchange itself. First-order rate constants such as k_{off} and pseudo first-order rate constant $k_{on}' \equiv k_{on}[X]_{free}$ are accurately determined by dynamic NMR techniques[19].

$$K_D = \frac{[P][X]}{[PX]} = \frac{k_{off}}{k_{on}} \qquad \text{(Eq. 1.2)}$$

$$k_{on} = \frac{k_{off}}{K_D} \qquad \text{(Eq. 1.3)}$$

According to previous literature (Table 1.1), the equilibrium dissociation constant for this interaction, K_D, is much lower than the concentrations of protein and peptide required for the NMR sample, which are in the order of millimolar, i.e. $[P]_{total} > [X]_{total} >> K_D$. As a result, the concentration of free peptide, $[X]_{free,}$ is vanishingly small in a partially saturated protein sample, on the order of nanomolar, and impractical to measure. The concentration of the bound state is governed only by the total concentration of peptide,

$[PX] \approx [X]_{total}$. In practice, even if k_{on}' can be experimentally determined, NMR-derived k_{off} values have to be complemented with equilibrium dissociation constants, K_D, measured by ITC, in order to determine k_{on} values (Eq. 1.2).

1.5 Identification of reaction intermediates

Where appropriate, more complicated reaction schemes have to be constructed in systems where experimental data[6,34,35] signal the presence of reaction intermediates or excited states. NMR is particularly well-suited to discriminate between two-state and multi-state binding kinetics since each nucleus in the protein and ligand whose chemical shift changes upon complex formation represents an independent probe of the binding reaction. The existence of binding intermediates may be inferred when data for multiple nuclei are incompatible with a global two-state mechanism[20]. In comparison, only a single observable is monitored in a typical non-equilibrium mixing experiment, yielding data sets that are insufficient to allow this sort of consistency tests. Inferences on the characteristics of reaction intermediates, if any, are more robustly supported by multiple independent probes of the binding reaction.

1.6 Characterization of the transition state

The Eyring–Polanyi equation (Eq. 1.4) relates the temperature-dependence of a reaction rate to the Gibbs free energy of activation $\Delta G^{\ddagger}$ for this reaction, where k_B is the Boltzmann constant, T is the temperature, h is Planck's constant, and R is the gas constant. This analysis is applied both to association and dissociation reactions and yields forward and reverse activation free energy values.

$$k = \frac{k_B T}{h} \cdot \exp\left\{\frac{-\Delta G^{\ddagger}}{RT}\right\} \qquad \text{(Eq. 1.4)}$$

The equation is rearranged to construct the Eyring plot (Eq. 1.5), with axes of $\ln\{k/\mathrm{T}\}$ versus 1/T, allowing the enthalpy $\Delta H^{\ddagger}$ and entropy $\Delta S^{\ddagger}$ components of the transition state stability to be easily investigated. Such Eyring plots are presented in Figure 4.4.

$$\ln\left\{\frac{k}{\mathrm{T}}\right\} = \frac{-\Delta H^{\ddagger}}{\mathrm{R}} \cdot \frac{1}{\mathrm{T}} + \frac{\Delta S^{\ddagger}}{\mathrm{R}} + \ln\left\{\frac{\mathrm{k_B}}{\mathrm{h}}\right\} \qquad \text{(Eq. 1.5)}$$

The slope ($-\Delta H^{\ddagger}/R$) is related to the enthalpy of activation, while the intercept can be related to the entropy of the transition state.

The enthalpy and entropy terms are related to the difference in heat capacity between states A and B (Eq. 1.6), where $\Delta H_0^{\ddagger}$ and $\Delta S_0^{\ddagger}$ are enthalpy and entropy values at the reference temperature T_0.

$$\begin{aligned} \Delta H^{\ddagger} &= \Delta H_0^{\ddagger} + \Delta C_P\left(\mathrm{T} - \mathrm{T_0}\right) \\ \Delta S^{\ddagger} &= \Delta S_0^{\ddagger} + \Delta C_P \cdot \ln\frac{\mathrm{T}}{\mathrm{T_0}} \end{aligned} \qquad \text{(Eq. 1.6)}$$

A greater heat capacity in the transition state than in the initial state ($\Delta C_P > 0$) leads to a negative curvature in the Eyring plot, while a decreased heat capacity leads to a positive curvature.

Protein association is commonly described using the following kinetic scheme[21-26]

$$P + X \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} P:X \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} PX \qquad \text{(Eq. 1.7)}$$

where *P*:*X* refers to the transient encounter complex, which may be thought of as an ensemble of molecular arrangements present at the end-point of translational diffusion[27].

Under steady-state conditions, association and dissociation rates are given by:

$$k_{on} = \frac{k_1 k_2}{k_{-1} + k_2} \qquad k_{off} = \frac{k_{-2} k_{-1}}{k_{-1} + k_2} \qquad \text{(Eq. 1.8)}$$

This scheme can appropriately be used to describe a two-state binding reaction, as long as the free energy of the encounter complex is large compared to that of the bound state, i.e. $G_{P:X} >> G_{PX}$. In this case, the bound and/or free states are the only significantly populated species in solution.

k_1 is a second-order rate constant describing the frequency of "successful encounters", i.e. when two binding surfaces are brought into favourable contact through random collisions. Dissociation of the encounter complex occurs with a rate constant k_{-1}, while rearrangement of the encounter complex into the bound state occurs with rate constant k_2.

Diffusion limit

In the diffusion limit, the rearrangement step is much faster than the dissociation of the encounter complex, $k_2 >> k_{-1}$. In this regime, association rates k_{on} are governed by the frequency of productive collisions k_1 which depends on the total collision frequency, subject to electrostatic enhancement, and a temperature-independent geometric factor ρ_{geo}[21].

$$k_{on} \approx k_1 = \rho_{geo} \cdot k_{coll} \qquad \text{(Eq. 1.9)}$$

The dissociation rate is determined by the stability and dissociation rate of the encounter complex:

$$k_{off} = \frac{k_{-2}}{k_2} k_{-1} \qquad \text{(Eq. 1.10)}$$

Activated reaction

If there exists a significant energy barrier between the encounter complex and the bound complex, only a fraction of the P:X population has sufficient energy to proceed to the bound complex. In this situation, the rearrangement step is much slower than the dissociation of the encounter complex, $k_2 << k_{-1}$. The association rate effectively depends both on the rate of rearrangement and on the stability of the encounter complex (Eq. 1.11) with a temperature-dependence described by (Eq. 1.4), where $\Delta G^{\ddagger}$ is the Gibbs free energy of the barrier between P:X and PX. In (Eq. 1.11), k_1/k_{-1} refers to the equilibrium association constant for the encounter complex.

$$k_{on} = \frac{k_1}{k_{-1}} k_2 \qquad \text{(Eq. 1.11)}$$

The off-rate directly depends on the dissociation rate of the complex, $k_{off} = k_{-2}$.

1.7 Thesis outline

The present Introduction chapter stated the motivation for the current research project and introduced some general concepts of protein binding kinetics. Chapter 2 will presents our results on the measurement of binding kinetics and detection of binding intermediates. Chapter 3 is concerned with the measurement of binding thermodynamics. Chapter 4 describes the diffusion properties of our system, as well as its rate-limiting steps for association and dissociation. Chapter 5 summarizes our findings and discusses their general application and biological relevance. Finally, involved details on the experimental methods are provided in Appendix.

Table 1.1 List of ligands for the Fyn SH3 domain

Name	Sequence	Type	K_d (μM)	Tech.	Ref.
"I"c+f'	VSLAR**R**PL**P**PL**P**	1	0.43	Fluo.	[28]
"I"c+f	*Ac*VSLAR**R**PL**P**PL**P**GGK*NH₂*	1	0.6	Fluo.	[17]
"II"c+f	*Ac*KGGGAA**P**PL**P**P**R**NRPRL*NH₂*	2	1.7	Fluo.	[17]
ΔSH3p85	*645 amino acids*	2	3	ITC	[14]
Sap	*104 amino acids* (no PxxP)	—	3.45	ITC	[29]
p85α2	PP**R**PL**P**VA**P**	1	4.41	Fluo.	[22]
p85α2	**K**PR**P**PR**P**LPVA	1	10	Fluo.	[30]
Nef	T**P**PV**P**P**R**PM	2	12.3	Fluo.	[22]
P2L	PP**R**PL**P**VA**P**GSSKT	1	16	ITC	[20,27]
Sos1−4	HSIAG**P**PV**P**P**R**	2	20	Fluo.	[30]
"I"core	*Ac***R**PL**P**PL**P**GGG*NH₂*	1	25.5	Fluo.	[17]
3BP2	PPAYPPPPVP	??	34	Fluo.	[16,30]
3BP1	RAPTMPPPLPP	??	34	Fluo.	[30]
Sos1−3	PPESP**P**LL**P**P**R**	2	48	Fluo.	[30]
FAK3A	AAAA**R**AL**P**SI**P**KL	1	68	ITC	[31]
p85α1	NERQPAPALPPKG	??	83	Fluo.	[30]
"II"c	*Ac*KGGGAA**P**PL**P**P**R***NH₂*	2	144.4	Fluo.	[17]
Nef	PVR**P**QV**P**L**R**PPMT	2	202	ITC	[21]
p85p3	PP**R**PT**P**VA**P**GSSKT	1	300	NMR	[13]
p85p1	KKISPPTPKPRPPR	??	3000	NMR	[13]
Map2	SEKKVAII**R**TP**P**KS**P**AT	1	N/A	ELISA	[31]
Tau	EPKKVAVV**R**TP**P**KS**P**SS	1	N/A	ELISA	[31]

CHAPTER 2: MEASURING KINETICS OF ASSOCIATION-DISSOCIATION, DETECTING BINDING INTERMEDIATES

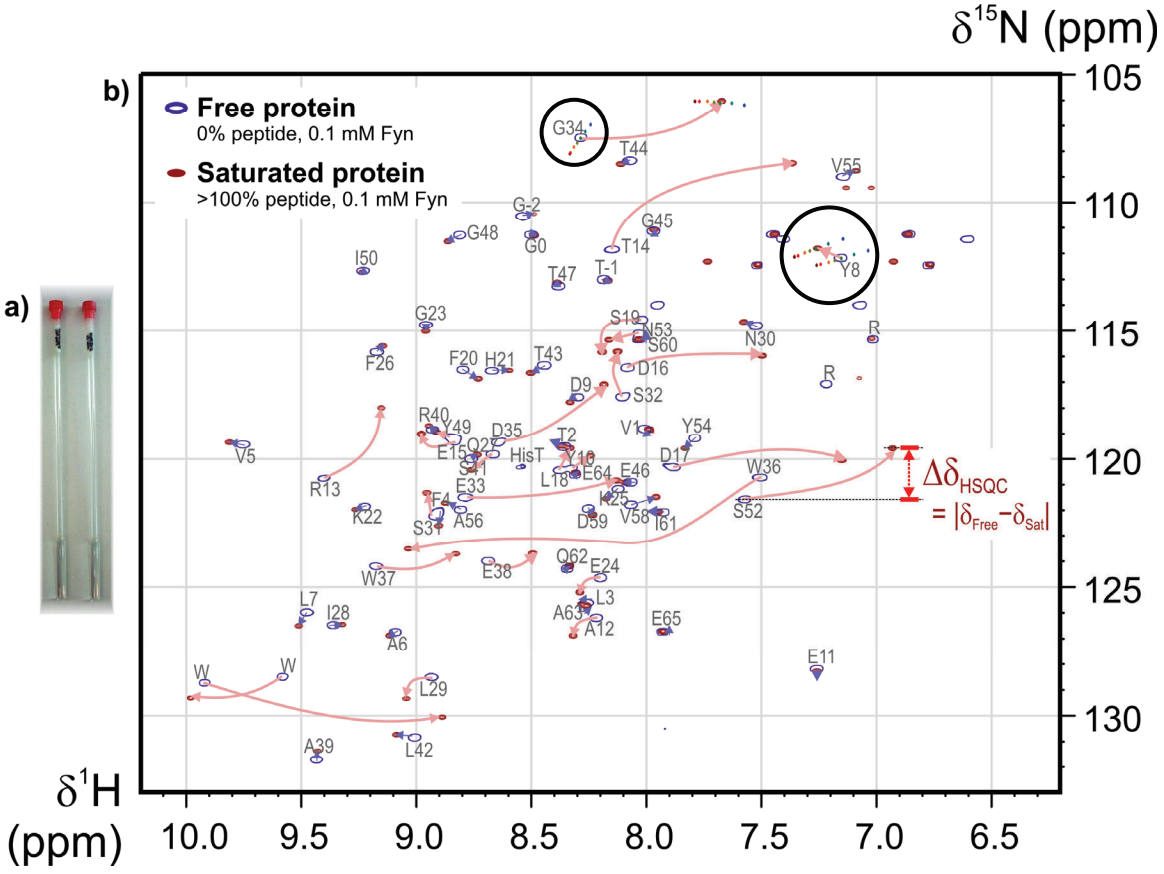
2.1 Analysis of chemical shifts

NMR resonances for all 66 backbone amide nuclei, with the exception of the 6 histidines from the affinity tag, could be assigned to their respective residue in the primary sequence of our Fyn SH3 construct (Fig. 2.1). Chemical shift values could be extracted for both the *apo* form and the peptide-saturated form of the SH3 domain (Fig. 2.2). The chemical shift difference between the free and peptide-saturated resonances could be computed for all temperatures studied. The residues with the largest ^{15}N chemical shift changes are located in the vicinity of the binding site for Type 1 ligands, as identified from previous structures[21,25]. Those residues, comprising R13, T14, S19, S32, G34, D35, W36 and S52, see chemical shift deviations in excess of 1.0 ppm upon saturation with peptide. However, binding causes subtle chemical shift perturbations throughout the protein, with an average deviation of about 0.5 ppm.

2.2 Temperature dependence of NMR spectra

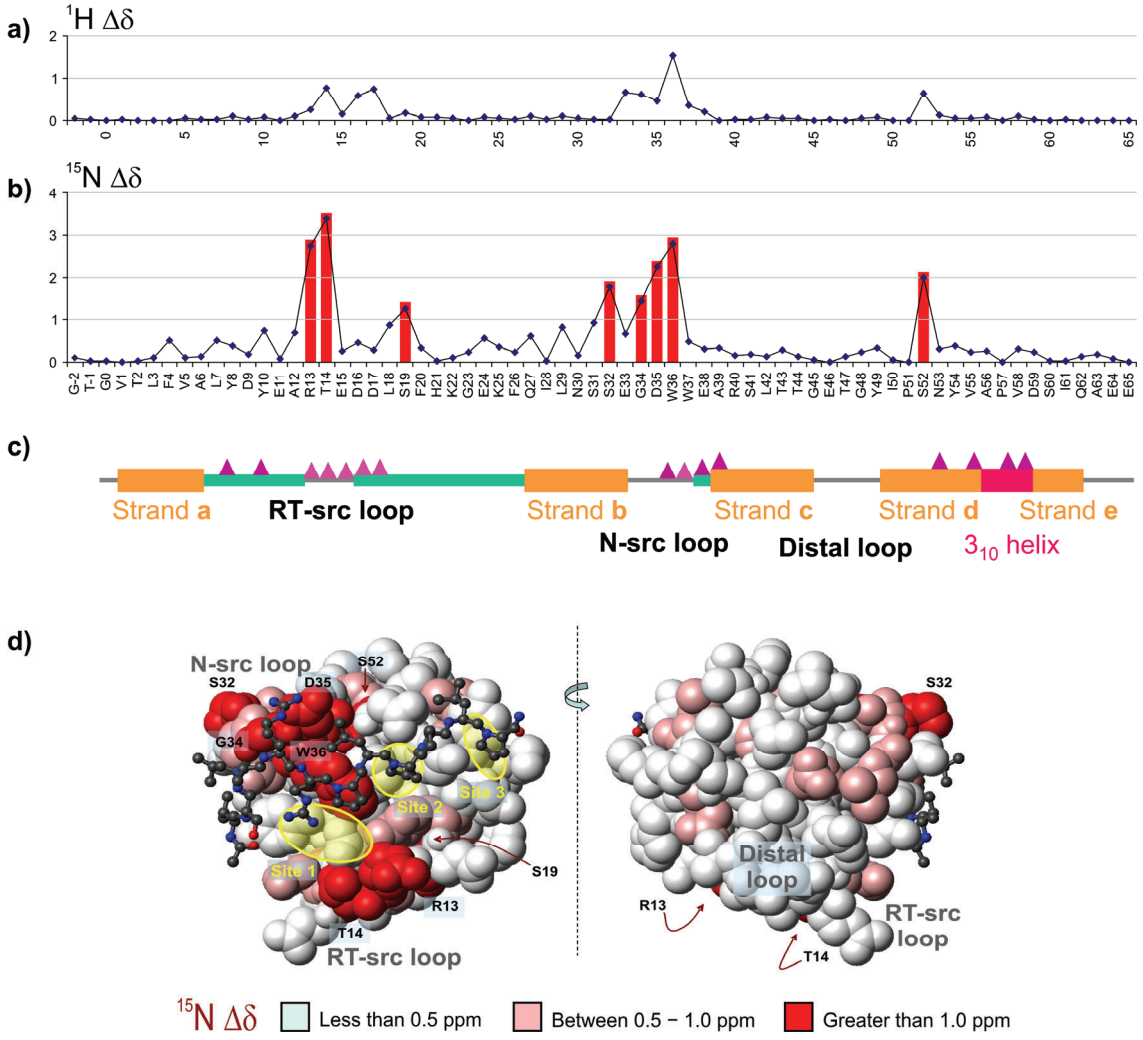
We employed a single NMR sample of ^{15}N-enriched Fyn SH3 domain that was partially saturated with a target peptide (Ac-VSLARRPLPPLP-NH_2)[17] for the binding kinetic experiments. At low temperatures (10°C and 20°C) , exchange was slow enough for separate bound and free protein peaks to be observed in ^{1}H-^{15}N correlation spectra. At higher temperatures (Fig. 2.1), the kinetics of exchange were more rapid. Amide resonances with large chemical shift differences (> 100 Hz) between the bound and free states were broadened beyond detection, while resonances with smaller chemical shift differences coalesced into single peaks, similarly to what is depicted in Figure 2.3.

Figure 2.1 HSQC spectra of the Fyn SH3 domain



(a) Two samples were prepared for resonance assignment, and determination of chemical shift deviations upon binding (see Section 6.2): first, the Fyn SH3 domain in its *apo* form, second, the Fyn SH3 domain saturated with peptide ligand. **(b)** The HSQC spectra from both samples, recorded at 30ºC, are superimposed, with free protein peaks in blue and bound peaks in red. The ^{15}N chemical shift deviation upon binding ($\Delta\delta_{HSQC}$) is measured from the vertical difference in peak positions between the free and bound peaks (red arrow). The amino acid assignments are indicated in grey for free peaks. Colourful dots for peaks G34 and Y8 (circled) give an example of the temperature dependence of the amide resonances, blue to red representing increasing temperatures from 10ºC to 50ºC.

Figure 2.2 Mapping of chemical shift deviations upon binding



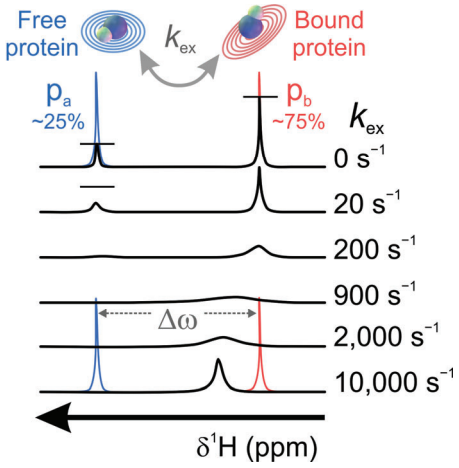
Amide ^{1}H **(a)** and ^{15}N **(b)** absolute chemical shift deviations upon binding (ppm) plotted along the primary sequence of the Fyn SH3 domain. For (b) ^{15}N $\Delta\delta > 1.0$ ppm are highlighted by red bars. (c) Secondary structure elements for the Fyn SH3 domain are plotted along the primary sequence, with element boundaries as defined in Larson *et al*, 2000[9]. Purple triangles identify conserved residues involved in binding, as identified by Larson. A high correlation is readily observable between conservation and magnitude of ^{1}H $\Delta\delta$ and ^{15}N $\Delta\delta$. **(d)** Mapping of ^{15}N $\Delta\delta$ on the structure of the Fyn SH3 domain. Residues with ^{15}N $\Delta\delta > 1.0$ ppm (red) are located in the vicinity of the binding site. Subtle chemical shift perturbations, $0.5 \text{ ppm} < \Delta\delta < 1.0$ ppm, are nevertheless seen throughout the protein

2.3 Conventional methods for kinetics measurement

Flow methods

In a typical non-NMR binding experiment, the reaction is initiated by the rapid addition of ligand, and the progress of the reaction is detected in real time, using a variety of measures which, in most cases, can be linearly related to the concentration of one or more of the species in solution. Those measures include fluorescence[4,32,33], spectroscopic absorbance[34], surface plasmon resonance (also discussed in Section 4.5)[35,36], or enzymatic activity[37-39]. The change of concentration with respect to time is then interpreted with a rate law to be able to extract rate constants. The lifetimes of the free and bound states must be several-fold longer than the time required to completely mix the protein and ligand solutions, otherwise the reaction proceeds nearly to completion before the start of the measurement period.

Figure 2.3 1D spectra for a two-site exchange system



Red and blue peaks correspond to spectra of states "a" and "b" as they could be acquired in isolation. Black spectra correspond to the simulated ^{1}H spectrum for a mixture of states "a" and "b" in a two-site exchange system, with the indicated rate of exchange and a population $p_b = 75\%$. Individual peaks can be observed for "a" and "b" states in the slow exchange regime. A frequency-averaged peak is observed in the fast exchange regime. Exchange causes preferential broadening of the minor population "a". Adapted from Palmer *et al*, 2001[1].

Line-shape analysis

NMR line-shape analysis has been used for many years to determine the rates of reversible chemical reactions in the ms-μs timescale[21]. In a nutshell, theoretical NMR

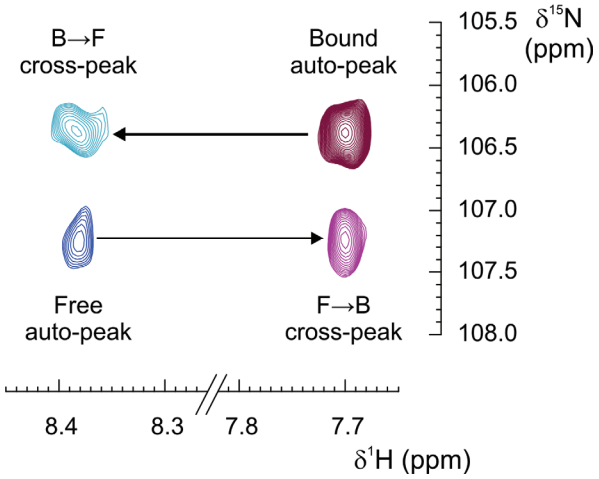
spectra can be predicted from a given kinetic exchange scheme, using either modified Bloch equations or the Density Matrix method[19] (see Fig. 2.3). Exchange parameters can then be extracted by finding the best fit between predicted and experimental spectra. Line-shape analysis is however ill-suited to study nuclei with low sensitivity, such as ^{15}N and ^{13}C, which are usually observed in indirect dimensions in multi-dimension NMR sequences. The limited number of data points does not allow for sufficient resolution to apply line-shape analysis in the case of those nuclei[1].

2.4 Magnetization exchange method (ZZ)

Principle

ZZ-magnetization exchange NMR experiments are essentially inverse-detection ^{1}H-^{15}N correlation experiments, analogous to a ^{1}H-^{15}N HSQC. However, an adjustable delay of duration T_{mix} is inserted between the ^{15}N chemical shift encoding period (t_1) and the direct acquisition of the ^{1}H signal. Protein amide magnetization is stored along the longitudinal ^{15}N axis during this period[40]. Concurrently, the exchange process produces novel NMR signals, allowing the measurement of association and dissociation kinetics.

Figure 2.4 ZZ exchange spectrum

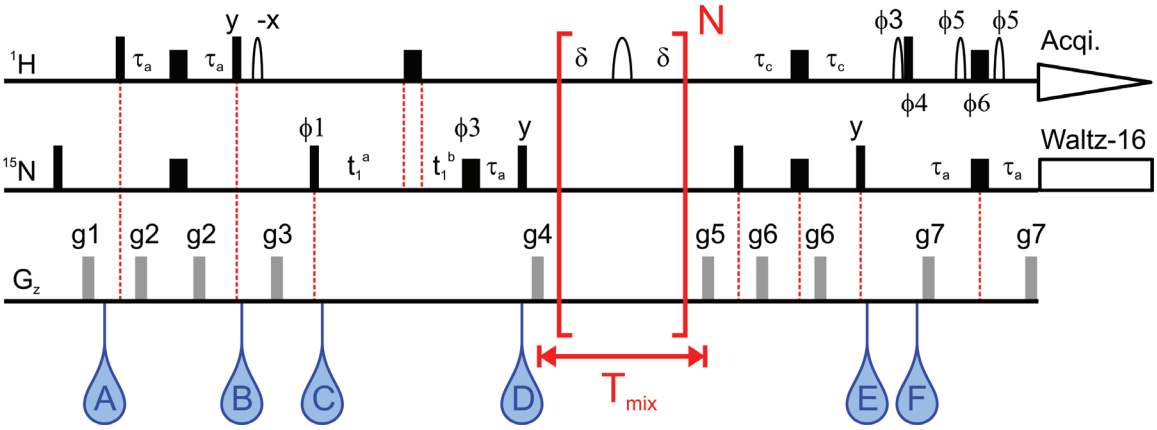


This spectrum excerpt demonstrates the appearance of cross-peaks in a ZZ exchange experiment for a set of tryptophane resonances at 10ºC. The spectrum was recorded with a mixing time of 0.3425 s. Only free and bound auto-peaks are visible in the HSQC spectrum of the partially-bound sample ($p_b \approx 74\%$). Stochastic binding exchange events during mixing time give rise to cross-peaks which conserve their initial ^{15}N frequency encoding but adopt a new frequency in the ^{1}H dimension.

For instance, a protein molecule which stays in the bound state throughout the pulse sequence will give rise to a resonance signal termed an 'auto-peak' that bears the ^{15}N and ^{1}H frequencies of the bound state. However, a free protein that binds peptide during T_{mix} gives rise to a cross-peak bearing the ^{15}N frequency of the free state but the ^{1}H frequency of the bound state. In the case of a two-state system, each amide group in the protein can produce four peaks: two auto-peaks corresponding to the free and bound forms, and two cross-peaks that are produced by peptide association and dissociation events (Fig. 2.4).

Multi-state systems give rise to additional auto- and cross-peaks. Rate constants associated with intermediate states can be extracted from the intensities of auto- and cross-peaks.

Figure 2.5 Pulse sequence for the ZZ exchange experiment



Pulse scheme for the ZZ exchange experiment. The following pulses are represented: hard 90º pulse (▮), 180º pulse (▮), selective pulse (▮), *z*-axis gradient pulse (▮). Transverse ^{15}N magnetization is generated at **C** by INEPT transfer. The signal is encoded in the ^{15}N dimension between **C** and **D**. Magnetization is placed along the *z* axis during the mixing time, allowing chemical exchange events to be observed. Magnetization is returned to the ^{1}H nucleus for acquisition. Adapted from Farrow *et al*, 1994[40].

Pulse sequence

Our treatment of peak intensities to extract rate constants differs from the standard method of Farrow *et all*, 1994[40]. To help appreciate the differences in analysis, the state of magnetization at various times in the pulse is briefly described. Highlighted letters in the text correspond to those particular time points, which are highlighted as well in Fig. 2.5. Please refer to Farrow *et all*, 1994[40] for a complete description of the pulse sequence.

The initial 90º pulse on ^{15}N and the gradient pulse dispose of initial $\hat{S}_z$ magnetization, such that only $\hat{I}_z$ magnetization is present (**A**). Magnetization is transferred with an INEPT block, such that $\hat{I}_z\hat{S}_z$ magnetization is generated at (**B**) and transverse ^{15}N magnetization at (**C**). Signal is encoded during the t_1 period according to the ^{15}N frequency. By (**D**), anti-phase magnetization is refocused. The 90ºy pulse on ^{15}N at that point rotates ^{15}N magnetization to the *z* axis where it will stay for a total duration of T_{mix}. Following the mixing period, magnetization is returned to $\hat{I}_z\hat{S}_z$ (**E**) through an INEPT block. Transverse proton magnetization is generated (**F**) for subsequent detection.

Measurement of rates

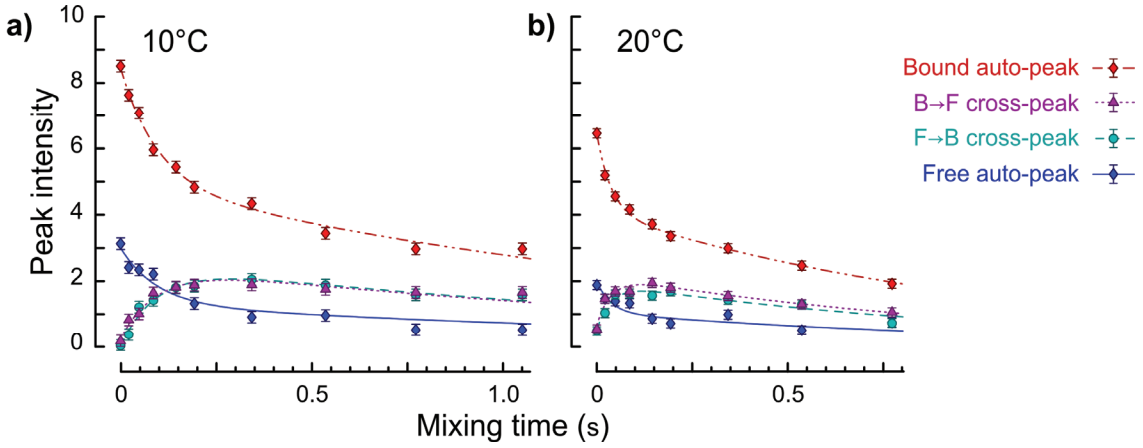
At 10ºC and 20ºC, separate peaks for the free and bound forms of the Fyn SH3 domain could be observed (Fig. 2.4). We restricted our analysis to residues for which both auto- and both cross-peaks are clearly resolved, which comprised 7 residues at 10ºC and 12 residues at 20ºC.

The intensities of auto-peaks and cross-peaks vary as a function of the mixing time T_{mix} (see Fig. 2.6) and depend on the dissociation rate, k_{off}, the fractional population of the bound state, p_B (Eq. 2.1), and the ^{15}N longitudinal relaxation rates.

$$p_B = \frac{k_{on}[X]}{k_{off} + k_{on}[X]} \qquad \text{(Eq. 2.1)}$$

At short mixing times, the experiment is characterized by the rapid build-up of cross-peak intensities due to chemical exchange, and by a corresponding decrease in auto-peak intensities. At longer mixing times, all peaks see their intensity decrease due to the dominating effect of longitudinal relaxation (Fig. 2.6).

Figure 2.6 Group fitting of ZZ exchange peak intensities



ZZ-exchange peak intensity profiles obtained for Gly34 of the Fyn SH3 domain in a partially ligand-saturated sample at **(a)** 10°C and **(b)** 20°C. Red (upper) and blue (lower) diamonds correspond to bound and free auto peak intensities. Teal circles and purple triangles correspond to association and dissociation cross peak intensities. Fitted curves were calculated using (Eq. 6.2) and globally-optimized dissociation rate constants and bound state populations.

In order to extract dynamic parameters, ZZ exchange data are usually fit to analytical equations that neglect exchange during magnetization transfer steps[1,40]. This is a reasonable assumption for slowly-exchanging systems, however in this case, the binding

kinetics are rapid and significant exchange occurs during the magnetization transfer steps (Fig. 2.5). This is particularly evident in the appearance of cross-peaks with the delay period, T_{mix}, set to zero (Fig. 2.6).

Data were fit numerically to an equation that accounts for exchange during the magnetization transfer steps, as described in section 6.6 (Eq. 6.2).

Group fitting

The values of k_{off} and p_B were optimized twice, once on a per-residue basis, and again as global parameters for all residues at a given temperature.

The group fitting increased the residual sum of squares by factors of only 1.08 (10ºC) and 1.20 (20ºC), compared to the individual fits (Table 2.1), and excellent agreement between the experimental and back-calculated intensities was obtained, shown in Figure 2.6. Dissociation rate constants extracted from the group fits are listed in Table 4.1.

2.5 Relaxation-based methods

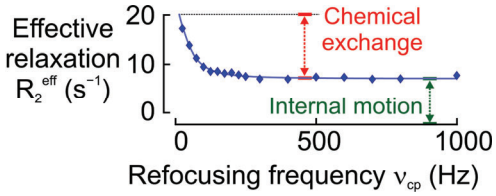
Figure 2.3 presents idealized 1D spectra of a two-site exchange system[1] with a skewed population ($p_a = 25\%$, $p_b = 75\%$) and various rates of exchange, k_{ex} (Eq. 2.2). As the exchange rate increases, the two resonances behave according to different spectral regimes depending on the relative magnitude of the frequency differences between the two sites, $\Delta\omega$ (Eq. 2.3).

$$k_{ex} = \left(k_{off} + k_{on}\left[X\right]\right) = \frac{k_{off}}{1 - p_B} \qquad \text{(Eq. 2.2)}$$

Slow exchange	$\Delta\omega << k_{ex}$	
Intermediate exchange	$\Delta\omega \approx 2k_{ex}\sqrt{p_a p_b}$	(Eq. 2.3)
Fast exchange	$\Delta\omega >> k_{ex}$	

Resonances for individual sites can be observed in the slow exchange regime; however, in the fast exchange regime, the observed peak is found at the population-weighted average frequency. Stochastic exchange events cause protein ^{15}N nuclei to experience different chemical environments, in which they do not precess at the same frequency. Fluctuations between states on the ms-µs timescale lead to the de-phasing of transverse magnetization, contributing to the R_2 relaxation rates of NMR signals. This contribution to relaxation is designated R_{ex}.

Figure 2.7 Contribution of chemical exchange to relaxation



Rapid conformational fluctuations in the ms-µs timescale contribute to enhanced transverse relaxation of NMR signals. By applying a train of spin echoes, with a repetition frequency ν_{cp}, it is possible to reduce transverse relaxation due to chemical exchange. At high ν_{cp}, the contribution of chemical exchange to transverse relaxation is practically eliminated, and the effective relaxation rates are similar to the relaxation term due to internal motion.

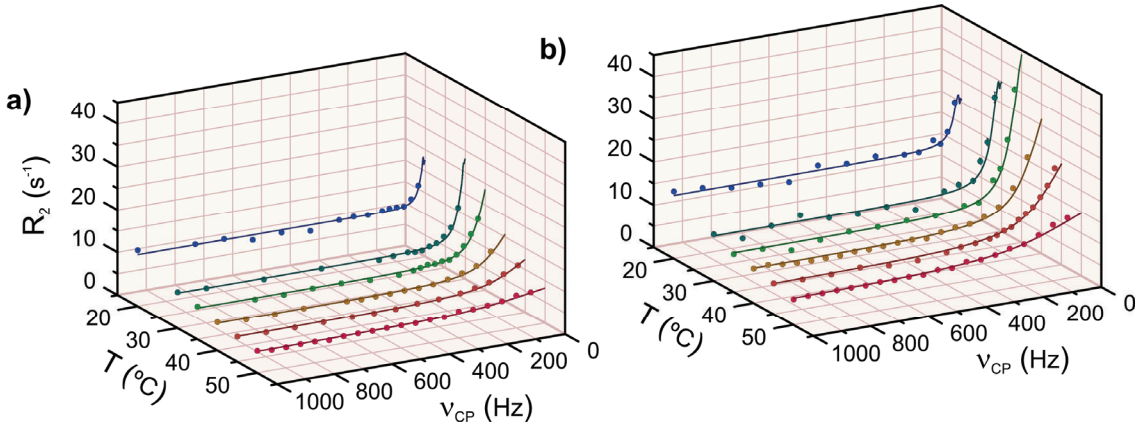
Two classes of experiment, CPMG and $R_{1\rho}$ experiments, explore the dependence of apparent relaxation rates on certain parameters to deduce the kinetics of chemical exchange between different environments.

In $R_{1\rho}$ experiments, radio-frequency fields are applied to spin-lock magnetization in the rotating frame. The strength of the spin-lock field is varied and the contribution of chemical exchange to rotating-frame relaxation ($R_{1\rho}$) changes accordingly. Although conceptually similar to CPMG, $R_{1\rho}$ can probe slightly faster time-scale motions (see Fig. 1.2), but it is less well suited to slower, ms-timescale processes. Hence, CPMG dispersion-relaxation experiments were used in this study.

CPMG experiments suppress the contribution of chemical exchange to ^{15}N transverse relaxation (R_2) using trains of 180° refocusing pulses[1] (Fig. 2.7). Assuming that the chemical shift fluctuations are caused by association and dissociation events as depicted in (Eq 1.1), R_2 will vary according to the delay between refocusing pulses (τ_{cp}) in a manner that depends on p_B, k_{off}, and the difference in precession frequency between the free and bound states, $\Delta\omega$.

Backbone ^{15}N CPMG relaxation dispersion experiments[41] were used to quantify the kinetics of Fyn SH3 domain binding at 20, 30, 35, 40, 45, and 50°C. Analytical expressions have been derived for fitting relaxation dispersion profiles of R_2 versus $\nu_{CPMG}=1/(2\tau_{cp})$ in various kinetic regimes[42-44].

Figure 2.8 Temperature dependence of CPMG curves



Plotting dispersion profiles, obtained at **(a)** 500 MHz and **(b)** 800 MHz, from Fig. 2.9 in a 3-dimensional perspective, shows how the aspect of the curves change with temperature.

Under intermediate exchange conditions, where $k_{ex} \leq \Delta\omega$, values of k_{off}, p_B, and $\Delta\omega$ may be extracted directly from the fits of relaxation dispersion profiles (as in Fig. 2.11a).

Under fast exchange conditions, where $k_{ex}>\Delta\omega$, the values of p_B and $\Delta\omega$ become co-dependent and cannot be extracted separately from the fits[1]. Instead, the parameters k_{ex} and Φ_{ex} are obtained (as in Fig. 2.11b), where

$$\Phi_{ex} = p_B(1 - pB)\Delta\omega^2 \qquad \text{(Eq. 2.4)}$$

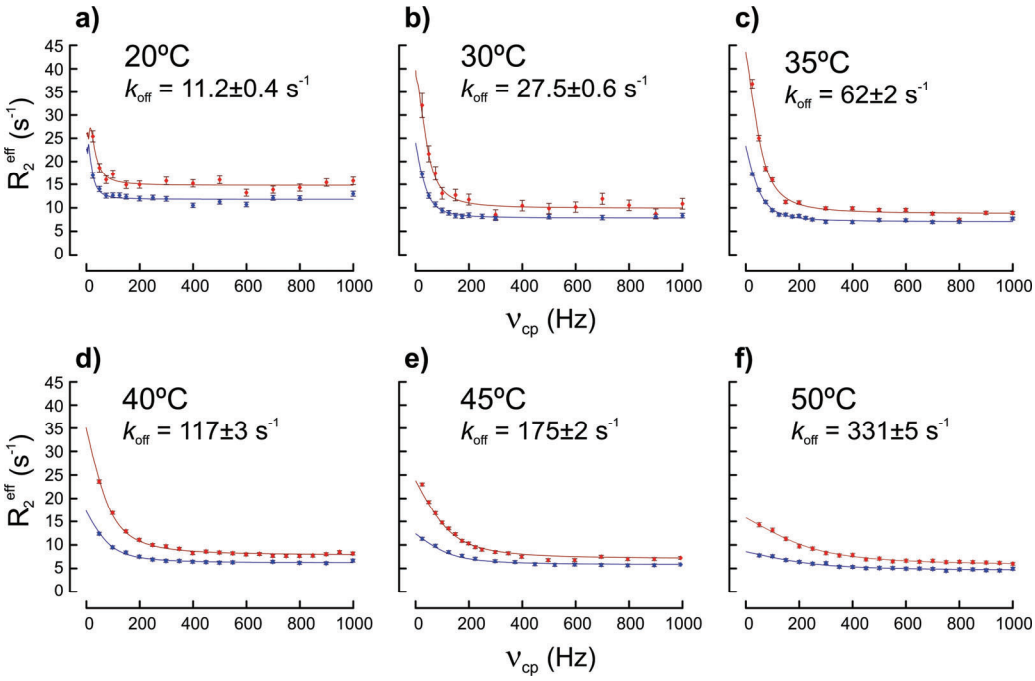
We fit the CPMG data by solving modified Bloch-McConnell equations numerically as described in section 6.7, using three models of increasing restriction:

- Model 1 consists of fits to individual amino acid residues. This model only serves as a benchmark for further fitting models.
- Model 2 tests for the presence of cooperativeness between residues by fitting k_{ex} and p_b as global variables.
- Model 3 fixes the values of chemical shift differences $\Delta\omega$ to those obtained from HSQC spectra, allowing k_{off} values to be extracted directly from residues in fast exchange, and with a higher precision in general.

Model 1: Fitting individually

For Model 1, each residue was fit independently at each temperature, yielding residue-specific values of p_B, k_{off}, and $\Delta\omega$ in the case of intermediate exchange, and residue-specific values of k_{ex} and Φ_{ex} in the case of fast exchange. This model does not correspond to a cooperative binding process since exchange at each residue is treated independently. The estimates of the exchange parameters obtained using Model 1 are of low precision, since the fits are highly unrestrained. We used the residual chi-squared values obtained from Model 1 as a benchmark for the goodness of fit provided by subsequent models (see Table 2.2) and did not further interpret the other parameters.

Figure 2.9 Group fitting of CPMG curves



^{15}N CPMG relaxation dispersion profiles obtained at 500 MHz (blue) and 800 MHz (red) proton Larmor frequencies for Leu7 of the Fyn SH3 domain in a partially ligand-saturated sample at 20, 30, 35, 40, 45, and 50°C. Best-fit curves were generated using Model 3, as described in the text, with fixed chemical shift differences and a temperature-independent value of the bound state population.

Model 2: Fitting as group

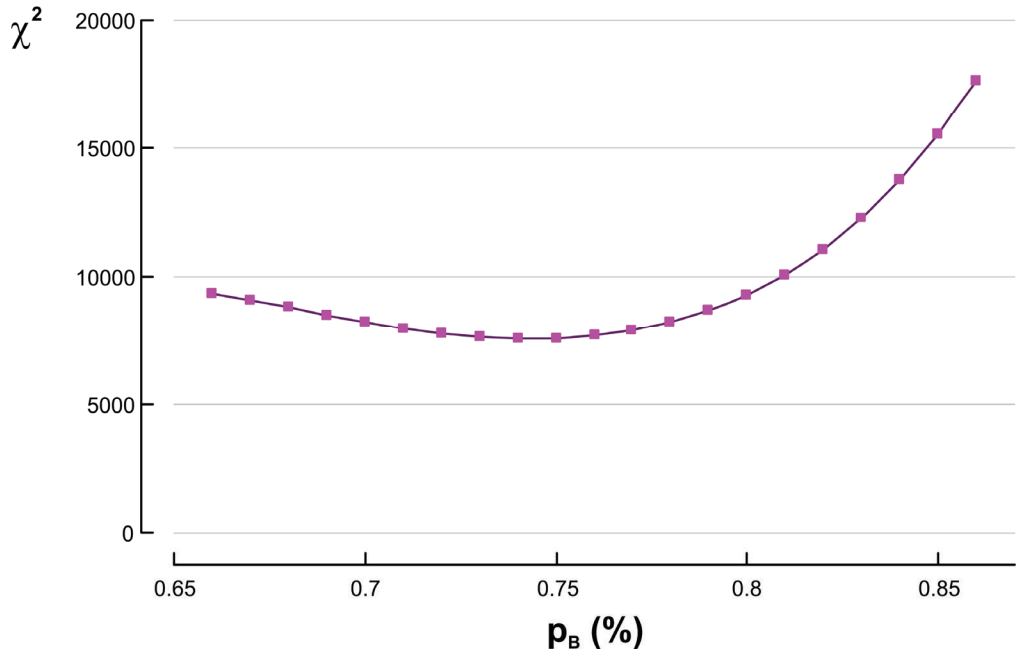
For Model 2, data at each temperature were analyzed assuming a cooperative two-state process, which increases the precision of the extracted parameters. For example, the uncertainties in global values of k_{ex} obtained with Model 2 are reduced by factors of 5 to 10 compared to those of Model 1. Such cooperative models have been successfully used to analyze CPMG data for a variety of different systems[4-6]. At 20°C and 30°C, exchange was in the intermediate regime for all residues; global values of p_B and k_{off} were extracted from the fits at each temperature, and Δω values were obtained on a per-residue basis (Fig. 2.9). At 35°C, 40°C, 45°C and 50°C, exchange shifted towards the fast regime due to the increase of the dissociation rate with increasing temperature, as illustrated by

shallower slopes in Fig. 2.8 and 2.9. The exchange rate k_{ex} was fit globally at each temperature, and values of Φ_{ex} were obtained on a per-residue basis. As discussed below, the good agreement of the extracted Φ_{ex} and $\Delta\omega$ values with HSQC-derived chemical shifts, as well as the similar magnitudes of residual chi-squared parameters obtained with Models 1 and 2 suggest that a mechanism of concerted two-state exchange between the free and bound states is sufficient to account for the CPMG data.

Model 3: Fixed chemical shift differences

For Model 3, additional restraints were imposed on the exchange equations in order to fit dissociation rate constants k_{off} with the highest possible level of precision. The values of $\Delta\omega$ were held fixed according to the differences in ^{15}N chemical shift between the free and peptide-saturated protein HSQC spectra. In addition, the population of the bound state, p_B, was constrained to be equal at all temperatures. This assumption is justified since the same sample was used in all cases, and the protein concentration was more than 100-fold greater than the equilibrium dissociation constant, K_D. Under these conditions, p_B is determined solely by the protein:peptide molar ratio, which is temperature-independent. Since the molar ratio is difficult to determine precisely by external measurements, we optimized this parameter using CPMG data and arrived at 0.74, as described in section 6.7 (see Fig. 2.10).

Figure 2.10 Optimisation of the bound protein population



In Model 3, optimization of the bound state population was accomplished by repeating fits with $\Delta\omega_{CPMG}$ set according to $\Delta\delta_{HSQC}$ values, and p_B fixed at a 0.01 intervals between 0.66 and 0.86. The the sum of χ^2 values over all temperatures shows a parabolic dependence on p_B with a minimum at p_B=0.74.

The residual χ^2 values obtained with Model 3 are about twice as large as those obtained from Models 1 and 2 (Table 2.2). This increase is not due to the constrained value of p_B, since allowing it to vary independently for each temperature while keeping the values of $\Delta\omega$ fixed does not significantly reduce the residual χ^2. The difference in χ^2 between Models 2 and 3 is therefore due to the fixed values of $\Delta\omega$. This is not surprising, since the agreement of CPMG- and HSQC-derived chemical shift differences is close, but not exact, due to uncertainties in both data sets. These discrepancies produce contributions to the residual χ^2 of Model 3 that are not present in those of Models 1 and 2.

Fixing $\Delta\omega$ values provides an advantage that outweighs this increase in χ^2. It resolves the co-dependence of $\Delta\omega$ and p_B values for residues in the fast exchange regime, so that

these data can be included in the global p_B optimization, and it allows k_{off} to be extracted directly from fits involving these residues. Despite the increase in χ^2, the agreement between experimental CPMG data and those back-calculated using Model 3 remains excellent, as shown in Figure 2.11. In addition, the value of k_{off} obtained at 20°C using Model 3 (11.2 s^{-1}) closely matches the ZZ-exchange-derived value (12.8 s^{-1}).

2.6 Two-State versus Multi-State Kinetics

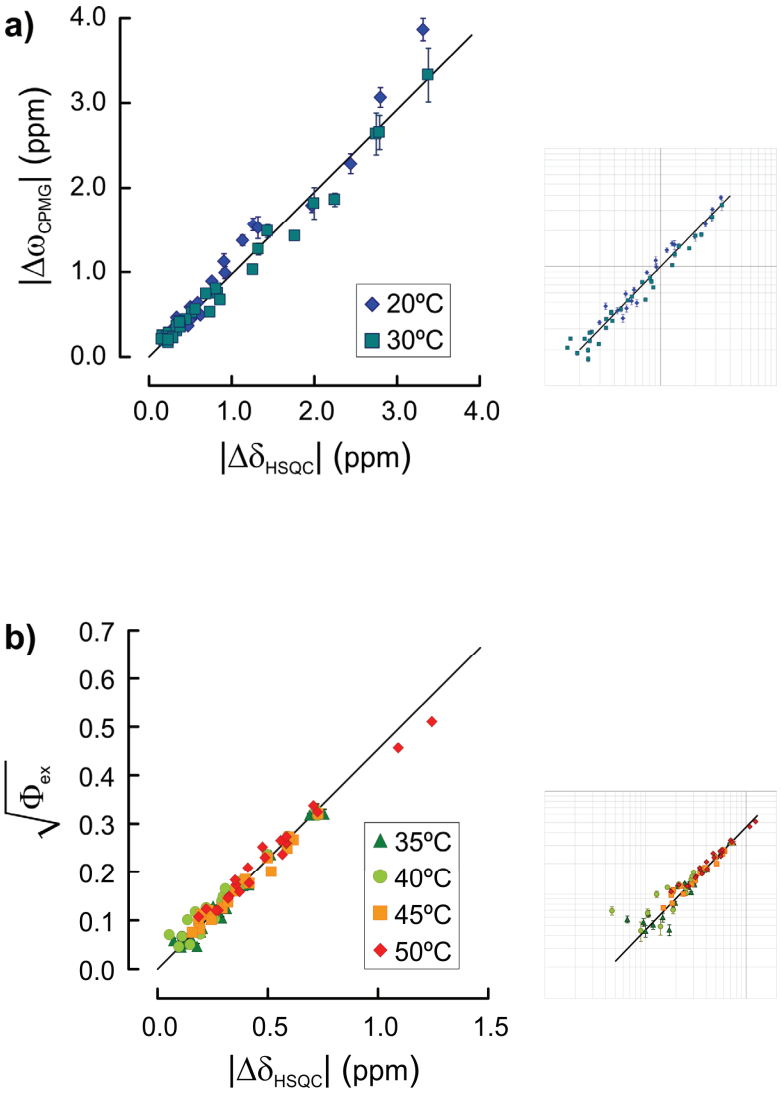
We tested the applicability of a two-state exchange model to the binding reaction in a number of different ways. First, we investigated the existence of pre-existing conformational equilibria in the SH3 domain by performing CPMG experiments on the free protein at 30°C and 50°C. These measurements produced flat dispersion curves for all residues, implying that ^{15}N chemical shifts do not experience significant fluctuations on the millisecond time scale in the absence of peptide. Therefore any alternative conformations of the free protein are either present at low populations (<0.5%), exchange extremely rapidly (>10^4 s^{-1}), or have identical ^{15}N chemical shifts to the major conformer.

We repeated these experiments at the same temperatures for a protein sample saturated with peptide at a 1:1.5 protein:peptide molar ratio, to test for the existence of partly-bound states in exchange with the fully-bound complex. The dispersion profiles were flat in this case as well, implying that millisecond timescale fluctuations of ^{15}N chemical shifts are absent in the complex. Thus any partly-bound states are either present at low populations (<0.5%), exchange with the bound complex extremely rapidly (>10^4 s^{-1}), or have identical ^{15}N chemical shifts to the bound complex.

A further piece of evidence that binding is effectively two-state is obtained from an inspection of the 2-dimensional ^{1}H-^{15}N NMR correlation spectra themselves. In all CPMG and ZZ-exchange spectra, as well as in a peptide titration performed at 30°C, the only observable signals correspond to the free, bound, and exchange peaks at low temperatures, or coalesced peaks at high temperatures. We did not detect any additional peaks corresponding to binding intermediates, as has been observed for the interaction of an SH2 domain with phosphotyrosine-containing peptides[45,46].

ZZ-exchange and CPMG data at all temperatures are consistent with a two-state binding model. The agreement of experimental peak intensities (Figure 2.6) and transverse relaxation rates R_2 (Fig. 2.8 and 2.9) with those back-calculated using global two-state exchange models is excellent. For ZZ-exchange data, residual sum of squares obtained using global exchange parameters are only 1.14-fold greater than those obtained with per-residue exchange parameters (Table 2.2). For CPMG data, residual χ^2 values obtained using the global exchange parameters of Model 2 are only 1.10-fold greater than those obtained with the per-residue exchange parameters of Model 1 (Table 2.2). In contrast, there are several examples of more complex systems that exchange among three or more states and produce NMR CPMG data that do not agree well with global two-state models[20,47,48].

Figure 2.11 Comparison of CPMG and HSQC chemical shift



(a) Chemical shift differences, $\Delta\omega_{CPMG}$, extracted from fits of ^{15}N CPMG data at 20°C and 30°C plotted as a function of the difference in ^{15}N chemical shift, $\Delta\delta_{HSQC}$, between the free and ligand-saturated forms of the Fyn SH3 domain. The line in (a) passes through the origin with a slope of 1. **(b)** Plot of $\sqrt{\Phi_{ex}} = \sqrt{p_B(1-p_B)}\Delta\omega$ values extracted from fits of ^{15}N CPMG data at 35, 40, 45, and 50°C versus chemical shift differences, $\Delta\delta_{HSQC}$. The line in (b) passes through the origin with a least-squares optimized slope. Insets figures are the double logarithmic plots of (a) and (b). They illustrate the robustness of correlations even at small chemical shift deviations.

To further test the two-state binding model, parameters $\Delta\omega$ or Φ_{ex}, extracted using Model 2 from CPMG fits, are compared with $|\Delta\delta_{HSQC}|$, the absolute differences in ^{15}N chemical shift measured in HSQC spectra of free and peptide-saturated protein. If the peak broadening observed in the partially-bound sample is due entirely to exchange between the free and bound states, a plot of CPMG- versus HSQC-derived parameters should be linear with a slope of 1. A plot for data obtained at 20°C and 30°C is shown in Figure 2.11a. At 35°C, 40°C, 45°C, and 50°C, the kinetics of the system enter the fast exchange regime, precluding the direct extraction of $\Delta\omega$ values. As discussed above, CPMG fits using Model 2 at these temperatures instead yield $\Phi_{ex}=p_B(1-p_B)\Delta\omega^2$. Again, if peak broadening is due to two-state exchange between free and bound states, then a plot of $\sqrt{\Phi_{ex}}$ versus $\Delta\delta_{HSQC}$ should be linear with a slope of $\sqrt{p_B(1-p_B)}$, as is observed in Figure 2.11b. Notably, scatter-plots of the data obtained at 35°C, 40º, 45ºC and 50°C all overlap in one line. Data sets from those temperatures are therefore consistent with the same value of p_B. Plots for intermediate and fast regimes all exhibits good agreement between CPMG- and HSQC-derived parameters, as expected for a two-state binding reaction. Taken together, these observations indicate that binding of the Fyn SH3 domain to the peptide used in this study is effectively two-state. Any binding intermediates are either weakly populated or in very rapid exchange with the free or bound states.

Table 2.1 Comparison of residual sum of square for ZZ exchange fits

10ºC

Residue	Indiv. RSS	Group RSS	RSS Ratio
Ser19	3.3261	3.6587	1.099997
Leu29	1.0331	1.0365	1.003291
Gly34	0.9573	0.9667	1.009819
Trp36	0.9574	0.9714	1.014623
Ser52	0.7839	1.0583	1.350045
Trp ε	0.5706	0.5808	1.017876
Trp ε	0.6565	0.6799	1.035644
Total	8.2849	8.9523	1.080556

20ºC

Residue	Indiv. RSS	Group RSS	RSS Ratio
Tyr8	0.725	1.205	1.662069
Arg13	0.8524	0.9222	1.081886
Asp17	0.7723	1.0279	1.330959
Leu29	0.5462	0.552	1.010619
Gly34	0.5062	0.5311	1.04919
Ser52	0.3956	0.4961	1.254044
Trp36	0.5656	0.5738	1.014498
Trp37	0.6581	0.7406	1.125361
Glu38	0.5874	0.8398	1.42969
Trp ε	0.2074	0.224	1.080039
Trp ε	0.2782	0.3022	1.086269
Asp35	1.101	1.2154	1.103906
Total	7.1954	8.6301	1.199391

Table 2.2 Comparison of 3 models to fit CPMG data

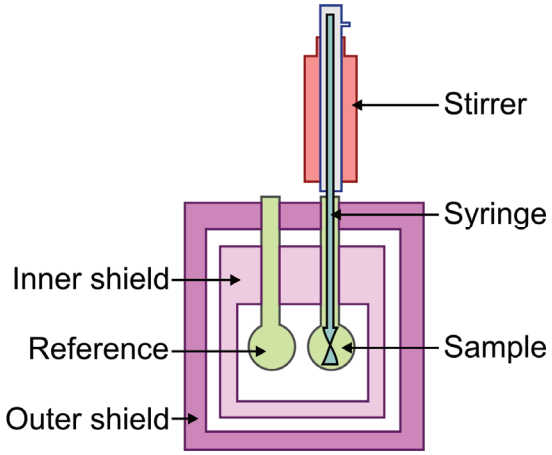
Temperature (ºC)	20	30	35	40	45	50
Number of probes	19	32	25	20	20	22
Cumulative χ^2						
Model 1	598.5404	474.4937	729.2003	565.6487	321.285	531.7635
Model 2	642.508	669.7289	695.564	586.7862	388.7432	559.7314
Model 3	804.372	1919.409	1855.767	1049.409	730.9981	1177.036
Degrees of Freedom						
Model 1	456	666	754	651	580	756
Model 2	493	798	699	671	599	776
Model 3	512	830	724	626	619	797
Reduced χ^2						
Model 1	1.312589	0.712453	0.967109	0.868892	0.55394	0.703391
Model 2	1.303262	0.839259	0.9577	0.908351	0.647266	0.721303
Model 3	1.571039	2.312541	2.563214	1.676372	1.180934	1.476833

CHAPTER 3: THERMODYNAMICS OF BINDING

3.1 Measuring binding thermodynamics with ITC

Isothermal titration calorimetry is a technique commonly used to measure the thermodynamic parameters of a binding reaction. The apparatus consists of two cells, the reference and the sample cells, whose temperatures are continuously monitored during the experiment (Fig. 3.1). Slight temperature differences, due to heat evolved or absorbed during the reaction, are compensated by a feedback mechanism and the power that is required to keep the two cells at the same temperature is recorded (Fig. 3.2a).

Figure 3.1 Overview of the ITC apparatus



The ITC apparatus consists of two cells, the reference and the sample cells, whose temperatures are continuously monitored during the experiment. Slight temperature differences are compensated by a feedback mechanism which controls the temperature of the shields. The ligand is injected from the injection syringe in small increments into the sample cell. To keep the sample uniform, the injection syringe is continuously stirring for the duration of the experiment.

The ligand is injected from the injection syringe in small increments into the sample cell. For each injection peak, the amount of power is integrated over time and normalized by the amount of ligand injected, correcting for dilution effects (Fig. 3.2b). For a two-state system, the initial heat relates to the enthalpy of the reaction. As binding sites are progressively filled with ligand, the heat evolved will be reduced until only the

background heat, related to the mixing process, is measured. The steepness of the isotherm is related to the affinity of the ligand for its binding partner.

In our case, the peptide ligand was titrated into a sample containing the Fyn SH3 domain. The thermodynamics of binding were characterized by ITC at 10ºC, 20ºC, 30ºC, 35ºC, 40ºC, and 50ºC. An appropriate protein concentration M_{tot} in the sample cell is selected based on the parameter c (Eq. 3.1), where n is the number of binding sites on the protein and K_D is the equilibrium binding constant.

$$c = \frac{M_{tot} \cdot n}{K_D} \qquad \text{(Eq. 3.1)}$$

A high value, $c > 500$, leads to steep isotherms which make extraction of K_D a difficult task. Low values, $c < 5$, on the contrary can lead to incomplete saturation of sites, rendering the determination of the binding enthalpy ΔH inaccurate.

Preliminary experiments indicated that a Fyn SH3 domain concentration of 20 μM would give a value for $c = 300$ to $c = 20$ between 20ºC and 50ºC. Further experiments confirmed that the correct range was selected, $c = 278$ to $c = 14$. Due to higher affinity at lower temperature, the protein stock had to be diluted to 4 μM for measurement at 10ºC, for a final value of $c = 107$.

As well, at least 10 separate injections are required to define the isotherm, with an average of 5 μcal for each injection, giving a concentration of $M_{tot} \geq 3.1\,\mu M$ since the lowest enthalpy recorded is $\Delta H_D = 11.2$ kcal/mol at 10ºC (Eq. 3.2).

$$M_{tot} \geq \frac{Q_{tot}}{V_{cell} \cdot |\Delta H|} \qquad \text{(Eq. 3.2)}$$

3.2 Fitting of ITC temperature curves

Good agreement was obtained by fitting the experimental isotherms to a simple two-state binding model, yielding the equilibrium dissociation constant, K_D, the enthalpy, ΔH_D, and entropy, ΔS_D, of dissociation at each temperature (Fig 3.2c-e). Values found for individual experiments are presented in Table 3.1. An exothermic binding reaction was observed at all temperatures, as illustrated in Figure 3.2a. The temperature dependences of these parameters are governed by the difference in heat capacity between the free and bound states, ΔC_p, according to (Eq. 3.3), (Eq. 3.4) and (Eq. 3.5), where R is the molar gas constant, and ΔH_0 and ΔS_0 are the enthalpy and entropy of dissociation at a reference temperature, T_0.

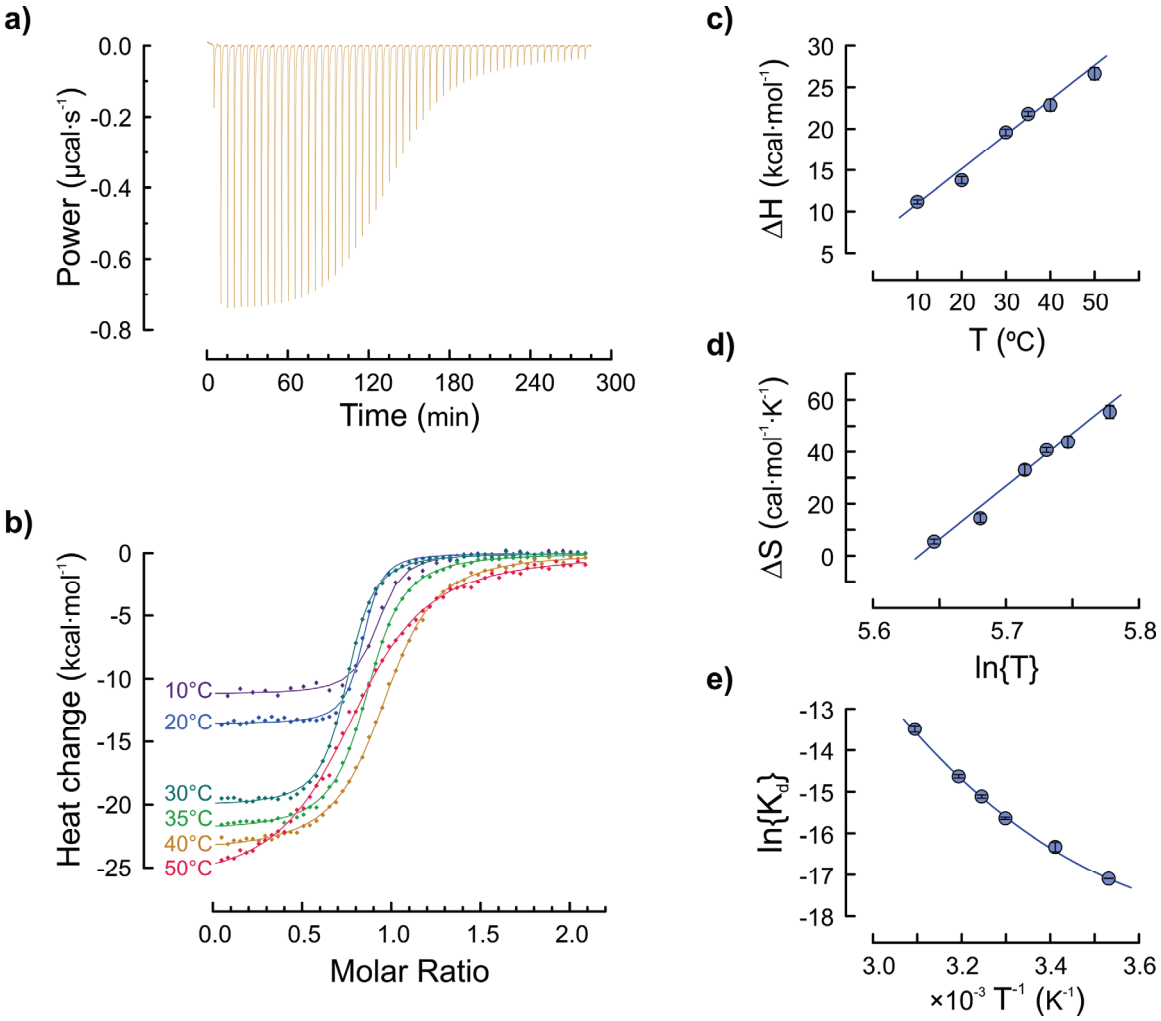
$$\Delta H_D = \Delta H_0 + \Delta C_p \left(T - T_0 \right) \qquad \text{(Eq. 3.3)}$$

$$\Delta S_D = \Delta S_0 + \Delta C_p \ln\left\{ \frac{T}{T_0} \right\} \qquad \text{(Eq. 3.4)}$$

$$K_D = \exp\left\{ \frac{-\Delta H_D + T\Delta S_D}{RT} \right\} \qquad \text{(Eq. 3.5)}$$

Plots of ΔH_D versus T and ΔS_D versus $\ln\{T\}$ (Fig. 3.2c-d) are both linear with slopes of ΔC_p, while a van't Hoff plot of $\ln\{K_D\}$ versus T^{-1} has curvature determined by ΔC_p (Fig. 3.2e). The dataset of ΔH_D, ΔS_D, and K_D values measured at 6 temperatures can be fit simultaneously to extract ΔH_0, ΔS_0, and ΔC_p.

Figure 3.2 Isotherms and fitting of the ITC data



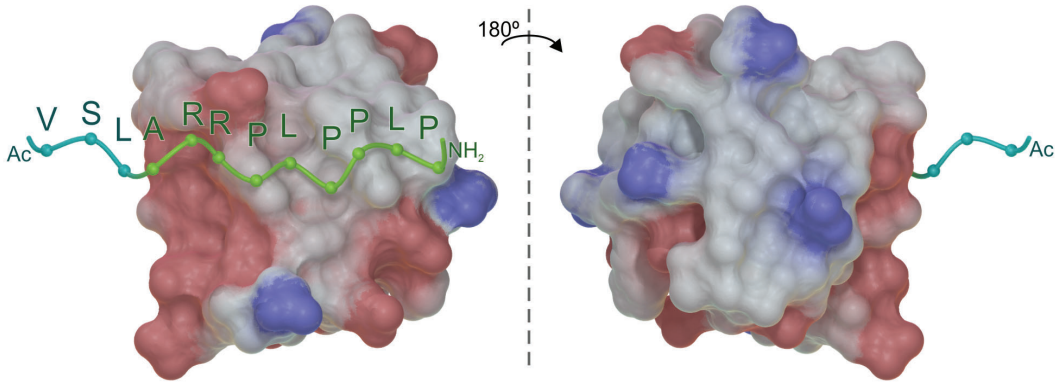
(a) Baseline-subtracted raw ITC data for injections of Ac-VSLARRPLPPLP-NH_2 peptide into a solution of the Fyn SH3 domain at 40°C. **(b)** Fits of ITC data to a simple two-state binding model. **(c)** Enthalpies of dissociation plotted as a function of reaction temperature. **(d)** Entropies of dissociation plotted as a function of $\ln\{T\}$. **(e)** van't Hoff plot of equilibrium dissociation constants. Lines and curve in (c)-(e) were generated with (Eq. 6.20) to (Eq. 6.22) using optimized values of ΔH_0, ΔS_0, ΔC_p, and β.

Correcting for concentration inaccuracies

When fitting the multiple-temperature ITC dataset, we found it necessary to introduce an additional correction factor, β, for the concentration of the injected peptide. The concentrations of the protein and peptide were thus adjusted separately in order to

optimize the agreement between experimental and calculated ITC data at all temperatures. Since the same peptide stock solution was used for all ITC measurements, this introduced only a single additional parameter to fits of the complete ITC dataset as described in section 6.9. An excellent agreement between experimental and back-calculated values was obtained with a correction value of β=0.88±0.02, with the standard deviation estimated by Monte Carlo simulation.

Figure 3.3 Polarity of the binding interface



Model of the protein-peptide complex. The Fyn SH3 domain is represented as a contact surface coloured according to the electrostatic potential, with positively charged regions in blue (>0.5 *e*/Å), neutral regions in grey, and negatively charged regions in red (<0.5 *e*/Å). The 12-mer peptide is represented by its C_α trace, with the positions of C_α atoms highlighted by spheres. Atomic positions for the peptide and the Fyn protein come from the PDB entry 1AZG, except for the first three C_α atoms of the peptide. The contact surface and electrostatic potential were calculated in Molmol and the image was rendered in POVray.

3.3 Peptide association is characterized by apolar desolvation

Taking the 30°C as the reference temperature, we found ΔH_0 = 15.4 ± 0.3 kcal mol^{-1}, ΔS_0 = 20.0 ± 0.9 cal·mol^{-1}K^{-1}, and ΔC_p = 370 ± 14 cal·mol^{-1}K^{-1}. These values are quite similar to those found in ITC studies of the Sem-5[49], Abl[50], and α-spectrin[51] SH3 domains. The large positive value of ΔC_p is consistent with hydrophobic nature of the binding interface, which includes the peptide residues Pro7 Leu8 Pro9 Pro10 Leu11 and

Pro12, adjacent to a hydrophobic patch on the SH3 domain surface composed of Tyr8, Tyr10, Trp36, Pro51, and Tyr54 (Fig. 3.3). Positive values of ΔC_p values in protein folding and binding are usually associated with the solvation of non-polar groups[52]. Dissociation of the peptide likely results in an increase in the solvent exposure of hydrophobic residues of both the protein and peptide and a concomitant increase in heat capacity.

Table 3.1 Experimental ITC values

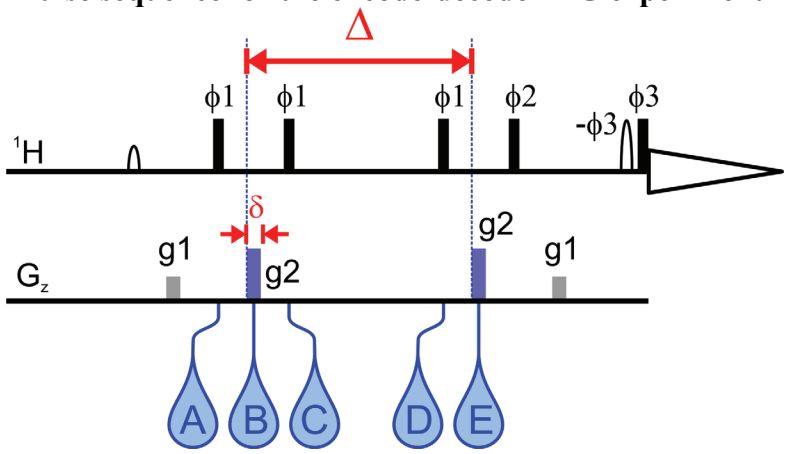
T (ºC)	10	20	30	35	40	50
K_D (nM)	37±4 38±2	76±8 69±4 93±29	185±19 151±6 164±9	270±22 291±7 252±6	453±21 357±22 449±10	1198±56 1196±40 1586±51
	37.5±0.2	72±5	158±10	272±19	442±28	1389±196
ΔH_D $\left(\frac{kcal}{mol}\right)$	10.58±0.09 11.38±0.05	13.26±0.09 13.58±0.05 15.15±0.3	17.95±0.2 20.00±0.06 19.16±0.08	19.83±0.2 21.76±0.05 22.03±0.05	21.04±0.1 20.97±0.1 23.69±0.06	24.8±0.2 26.4±0.2 28.2±0.2
	11.2±0.4	13.5±0.3	19.5±0.6	21.8±0.6	23±1	27±1
ΔS_D $\left(\frac{cal}{mol \cdot K}\right)$	6.2±0.5 3.4±0.7	19.5±2 12.7±0.7 13.5±0.5	28.4±0.8 32.1±0.5 34.8±0.4	40.7±0.4 34.3±0.8 41.3±0.4	38.2±0.6 46.6±0.4 37.5±0.7	60.7±0.6 49.7±0.8 54.7±0.6
	5±1	14±2	33±2	40±2	43±4	56±4

CHAPTER 4: CHARACTERIZING BINDING TRANSITION STATE

4.1 Measuring macromolecular diffusion properties: DOSY

The translational diffusion coefficients of the protein and peptide were determined using 1D 1H linear-encode-decode pulsed-field-gradient (LED-PFG) NMR experiments[53] at 10°C, 20°C, and 30°C. The pulse sequence of this experiment is briefly described.

Figure 4.1 Pulse sequence for the encode-decode PFG experiment



Pulse scheme for the LED-PFG exchange experiment. The following pulses are represented: hard 90º pulse (▮), selective pulse (▯), *z*-axis gradient pulse (▮). At **B**, transverse magnetization encoded with a spatially-dependent phase by a pulsed-field gradient of strength G_z and length δ. Between **C** and **D**, signal is stored for a constant diffusion delay as longitudinal magnetization. Signal is then decoded in a spatially-dependent manner by a second pulsed-field gradient **E** and eventually recorded. Adapted from Altieri *et al*, 1995[53].

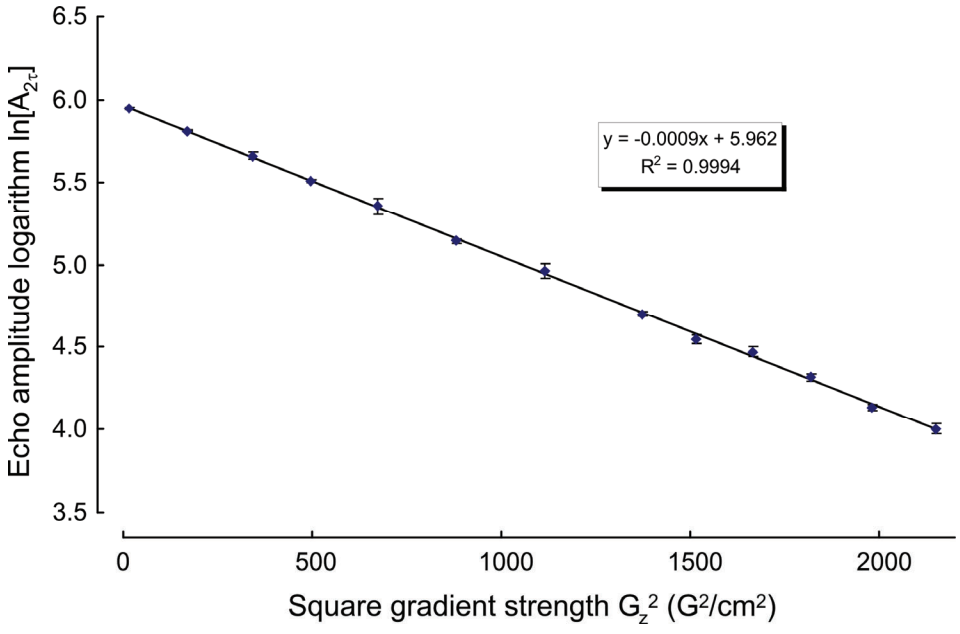
Transverse magnetization is first generated by a 90º pulse at point **A**. Magnetization is then encoded with a spatially-dependent phase by a pulsed-field gradient (**B**) of strength G_z and length δ. The signal is returned to the *z* axis (**C**) and stored as longitudinal magnetization for a constant diffusion delay.

After the delay, magnetization is brought back into the transverse plane (**D**). The phase is then decoded in a spatially-dependent manner by a second pulsed-field gradient (**E**).

Any distance that a macromolecule travels along the *z* axis, between time points **C** and **D**, will acquire a different phase from the *encode* and *decode* gradients. The amplitude of the echo generated by an ensemble of magnetically-equivalent diffusing nuclei can thus be related to a translational diffusion coefficient (see Appendix, section 6.8).

To ensure that only the decoded signal is observed, the signal is put on the *z* axis by the ϕ2 pulse, after which a *z*-gradient pulse eliminates residual transverse magnetization. The final ϕ3 pulse prepares transverse magnetization for acquisition, while the -ϕ3 selective pulse keeps water signal along the *z* axis.

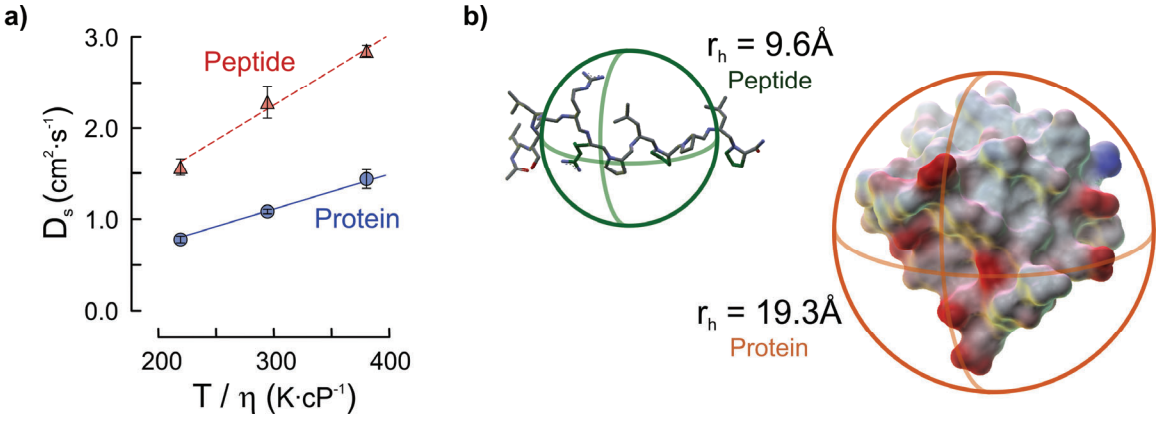
Figure 4.2 Extraction of the self-diffusion coefficient



LED-PFG experiments are repeated with a range of gradient strengths. The value of the diffusion constant is extracted from the slope of this graph according to Altieri *et al*, 1995[53].

Large gradient strengths, diffusion constants, and relaxation delays lead to large attenuation of the echo signal. The experiment is executed with a range of gradient strengths (Fig. 4.2). The value of the diffusion constant can then be extracted from the dependence of peak intensities on the strength of the gradients, G_z.

Figure 4.3 Protein and peptide diffusion properties



(a) Translational diffusion constants, D_s, for the Fyn SH3 domain (circles) and Ac-VSLARRPLPPLP-NH_2 peptide (triangles) determined by NMR plotted as a function of temperature/viscosity. **(b)** Spheres of hydrodynamic radius r_h, as determined by LED-PFG NMR experiment are superimposed on the molecular model of peptide and protein.

4.2 Protein and peptide diffusion follows Stokes-Einstein

The translational diffusion constants obtained for protein and peptide are plotted as a function of T/η (Fig. 4.3a), where η is the solvent viscosity. If the molecules obey the Stokes-Einstein diffusion equation (Eq. 4.1), the graph is linear with a slope equal to $k_B/(6\pi r_H)$, where k_B is Boltzmann's constant and r_H is the hydrodynamic radius. Data for both protein and peptide closely follow straight lines with correlation coefficients of greater than 0.99.

$$D = \frac{k_B T}{6\pi r_H \eta} \qquad \text{(Eq. 4.1)}$$

Hydrodynamic radii of 17.4 ± 0.4 Å and 9.2 ± 0.1 Å were obtained for the protein and peptide, respectively (Fig. 4.2b). The obtained radii agree with the estimated values predicted for molecules with this number of residues[54], as calibrated for unstructured poly-peptide chains, in the case of the peptide ligand, or globular proteins, in the case of the Fyn SH3 domain.

4.3 Transition State is solvated (Eyring plot of off-rates)

First-order dissociation rate constants k_{off}, as derived from longitudinal exchange (Section 2.4) and CPMG (Section 2.5) experiments, show a strong temperature dependence, varying from 4.4 s^{-1} at 10°C to 331 s^{-1} at 50°C (detailed values in Table 4.1). This large dependence implies that dissociation involves a significant enthalpic barrier.

We modeled the dissociation rate assuming that $k_2 >> k_{-1}$ and that k_{-1} follows Eyring kinetics[55]. Off-rates are given by (Eq. 4.2) to (Eq. 4.6), where G^{PX} is the free energy of the bound state, and G^{TS} is the free energy of the k_{-1} transition state. This state corresponds to the rate limiting step of the dissociation of the encounter complex as depicted in reaction scheme (Eq. 1.7). The terms $\Delta H^{\ddagger}$, $\Delta S^{\ddagger}$ and $\Delta C_p^{\ddagger}$ correspond to differences in enthalpy, entropy or heat capacity between the k_{-1} transition state and the bound state.

$$k_{off} = \frac{k_B T}{h} \exp\left\{\frac{-\Delta G^{\ddagger}}{RT}\right\} \qquad \text{(Eq. 4.2)}$$

$$\Delta G^{\ddagger} = G^{TS} - G^{PX} = \Delta H^{\ddagger} - T \cdot \Delta S^{\ddagger} \qquad \text{(Eq. 4.3)}$$

$$\Delta H^{\ddagger} = \Delta H_0^{\ddagger} + \Delta C_p^{\ddagger}\left(T - T_0\right) = H^{TS} - H^{PX} \qquad \text{(Eq. 4.4)}$$

$$\Delta S^{\ddagger} = \Delta S_0^{\ddagger} + \Delta C_p^{\ddagger} \ln\left\{T/T_0\right\} = S^{TS} - S^{PX} \qquad \text{(Eq. 4.5)}$$

$$\Delta C_p^{\ddagger} = C_p^{TS} - C_p^{PX} \qquad \text{(Eq. 4.6)}$$

Two hypotheses were tested with regards to the difference in heat capacity, $\Delta C_p^{\ddagger}$. A value of $\Delta C_p^{\ddagger} = 0$ gives a linear fit (red dotted line, Fig. 4.4b) of the experimental data on the Eyring plot of dissociation rates. In the second hypothesis, $\Delta C_p^{\ddagger}$ was fixed to match the ITC-derived value of $\Delta C_p = 370$ cal mol^{-1} K^{-1}. In that case, the fit (blue solid line, Fig. 4.4b) exhibits similar curvature to that of the van't Hoff plot in Figure 3.2e.

Curvature of the k_{off} Eyring plot is consistent with a hydrophobic protein/peptide interface that is solvated at the rate-limiting step of dissociation, such that the heat capacity of the transition state, C_p^{TS}, is greater than that of the bound state C_p^{PX}. Large changes in heat capacity are usually thought to accompany changes in molecular solvation[52]. The solvation of a non-polar solute, portrayed in Fig. 4.4c, is characterized by the formation of cage-like solvent structures at the solute-solvent interface. A linear fit of experimental off-rates, on the contrary, suggests that solvation occurs after k_{-1} in the dissociation process, leaving the TS and PX states with similar heat capacities.

The curved fit with $\Delta C_p^{\ddagger}$=370 cal mol^{-1} K^{-1} provides better agreement with the experimental data than (Fig. 4.4b), and the residual chi-squared is lower by a factor of 2.5. Further support is provided by the result of the Monte Carlo simulation (Fig. 4.4c). The χ^2 statistics, describing the lack of fit between two data sets, is higher for the linear fit with $\Delta C_p^{\ddagger}$=0 in systematically all cases (n=20,000) of the simulation. A jackknife procedure was carried in which a single temperature point was removed and both fits were repeated ($\Delta C_p^{\ddagger} = 0$ or 370 cal mol^{-1} K^{-1}). The procedure confirmed the importance of either the 10ºC or 30ºC data points for a significant difference in chi-squared ratio. The

original fits are reliable nonetheless, as the 10ºC and 30ºC points reflect the contribution of many independent probes of the dissociation reaction, respectively 7 and 32.

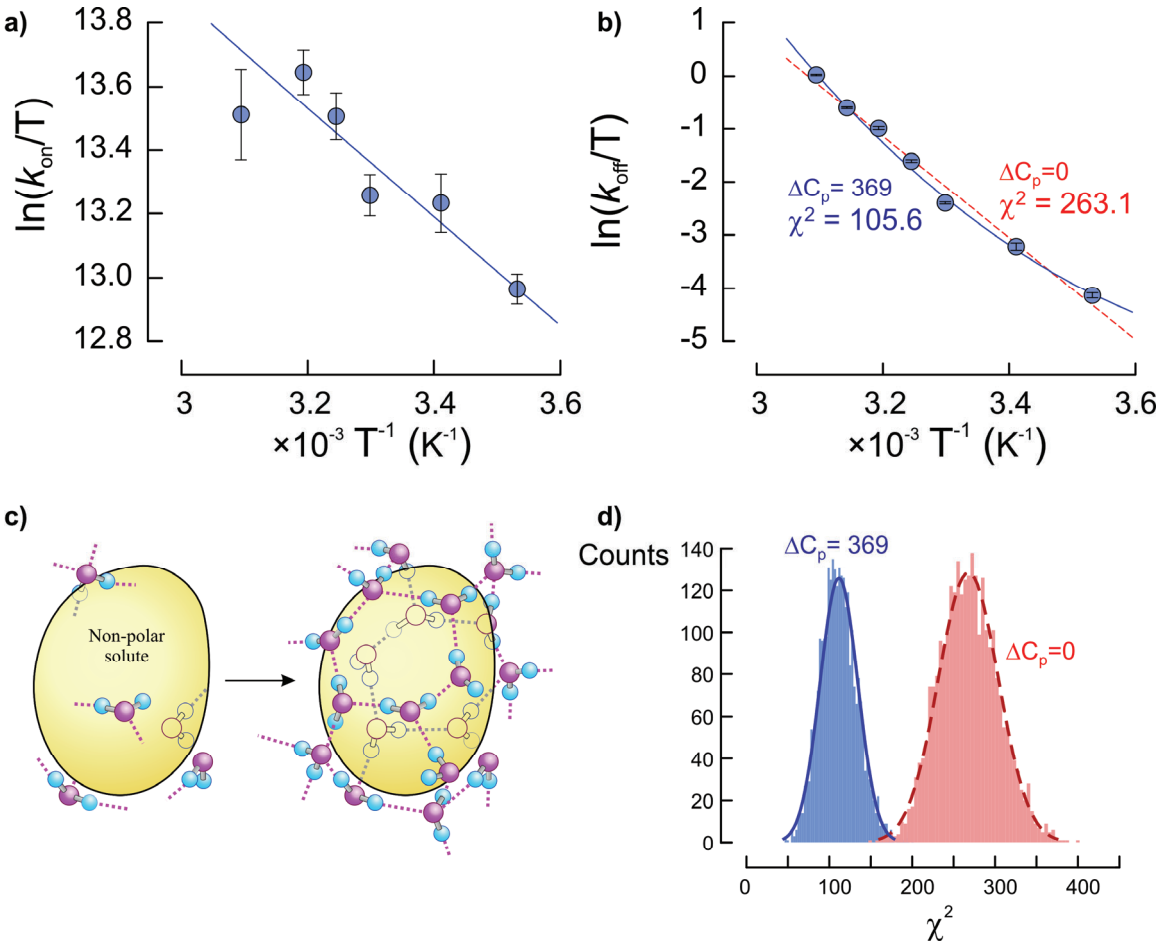
This curvature suggests that the hydrophobic portion of the protein/peptide interface is largely disrupted, and the non-polar groups it contains are hydrated during the rate-determining step of dissociation. In other words, re-solvation of the hydrophobic interface is a relatively early event in the dissociation process.

4.4 Association is diffusion-limited (Eyring plot of on-rates)

Second-order association rate constants, k_{on}, were calculated from the NMR-derived dissociation rate constants, k_{off}, and corrected equilibrium dissociation constants, K_D, following (Eq. 1.3). Values of k_{on} do not vary significantly with temperature (Fig. 4.4a), ranging from 1.1×10^8 M^{-1} s^{-1} at 10°C to 2.1×10^8 M^{-1} s^{-1} at 50°C, implying that association does not involve a large enthalpic barrier. Detailed values are found in Table 4.1.

An Eyring plot of $\ln\{k_{on}/T\}$ versus T^{-1}, shown in Figure 4.4a, has a slope of $-\Delta H_A^{\ddagger}/R$, where $\Delta H_A^{\ddagger}$ is the activation enthalpy of association. The level of agreement between the experimental points and the linear fit is quite encouraging, considering that the values of k_{on} vary by only a factor of 2 over the entire temperature range, and that they are obtained from a combination of ZZ-exchange and CPMG NMR kinetic experiments and ITC equilibrium measurements.

Figure 4.4 Eyring plots for association and dissociation



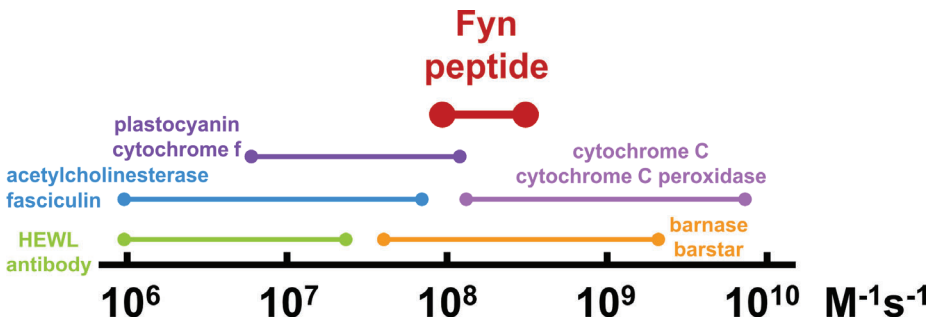
(a) Eyring plot of association rate constants calculated from ITC and NMR data according to (Eq. 1.3) **(b)** Eyring plot of NMR-derived dissociation rate constants. The solid curve and dashed line correspond to fits of the data with $\Delta C_p^{\ddagger}$=370 cal mol^{-1} K^{-1} and $\Delta C_p^{\ddagger}$=0, respectively, as described in the text. **(c)** Solvation of a non-polar solute. **(d)** Distribution of the χ^2 statistics generated by Monte Carlo analysis of the dissociation Eyring plot. Blue histograms correspond to linear fit with $\Delta C_p^{\ddagger}$=369, red histograms correspond to $\Delta C_p^{\ddagger}$=0. Contour lines represent the fit of normal distribution to those histogram plots. The overlap between the two distributions is minimal.

The activation enthalpy of association, $\Delta H_A^{\ddagger} = 3.4 \pm 0.4$ kcal mol^{-1}, obtained from the slope of the plot matches the expected value for a diffusion-limited binding reaction. This can be verified by replacing k_{on} values with the sum of NMR-derived diffusion constants in the Eyring plot, ($D_{Fyn}+D_{PxxP}$). The apparent activation energy obtained in this case, 4.4 ± 0.4 kcal mol^{-1}, is very similar to the actually enthalpy of association. The latter is also close to the characteristic value for a diffusion-limited reaction in water, approximately 4.5 kcal mol^{-1} [56]. Both the magnitude and temperature dependence of k_{on} values are thus consistent with a diffusion-limited binding reaction, suggesting that $k_2 >> k_{-1}$, i.e. that the folding and docking of the peptide and the rearrangement of the binding site occur rapidly compared to the dissociation of the encounter complex.

4.5 Comparison to other bimolecular association systems

The kinetics of association/dissociation in this system are much more rapid than those of other protein/ligand interactions for which the binding mechanism has been studied in detail. For example, the average survival time of the SH3 domain/peptide complex is just 3 ms at 50°C. This is about 7 orders of magnitude shorter than that of the well-characterized complex between barnase and its inhibitor barstar, which survives 35 hours on average before dissociation[4]. However, the association rates are similar to those of other protein-ligand complexes whose association is believed to be diffusion-limited[3], as illustrated in Figure 4.5. This rapid association with ligand could serve to allow a rapid activation of the kinase domain of Fyn (see Section 5.2).

Figure 4.5 Diffusion-controlled reactions



Comparison of the second-order association rates for Fyn:peptide to other classical bimolecular association systems.

Other groups have studied the kinetics of peptide binding by SH3 domains. Recently, Kay and co-workers used a suite of CPMG-based NMR experiments to characterize the interaction between the Abp1 SH3 domain and a fragment of the yeast Ark1p protein, obtaining a dissociation rate constant, k_{off}, of about 230 s^{-1} at 25°C[57]. Combining this number with a dissociation constant of 0.55 μM [58], the association rate constant, k_{on}, is 4.2×10^{8} M^{-1} s^{-1}, which agrees well with our results.

This contrasts with a number of SH3 domain studies performed using surface plasmon resonance (SPR), which have found much slower association and dissociation rate constants, on the order of 10^{3} to 10^{5} M^{-1} s^{-1} and 10^{-2} to 10^{-3} s^{-1}, respectively[59-63]. One explanation for this discrepancy is that the peptide/protein pairs chosen for SPR analysis simply happen to have much slower binding kinetics than the systems studied by NMR. Another possibility is that the systems studied by SPR do, in fact, follow similar kinetics to those observed by NMR, but that the SPR measurements are not reflective of solution-state association and dissociation rates. In SPR experiments, the mass transport of ligand into the hydrated matrix containing the immobilized binding partner, and impeded

diffusion of the ligand within this phase, can dominate the binding kinetics when the intrinsic reaction rate is greater than 10^6 M^{-1} s^{-1} [64]. This is well below the NMR-derived k_{on} values for SH3 domains. In such cases, the kinetic parameters determined by SPR can underestimate the true values by several orders of magnitude[65]. This underlines one of the advantages of equilibrium, solution-state NMR binding measurements. They do not rely on the net transfer of molecules from one phase to another and therefore do not suffer from mass transport artifacts.

Table 4.1 Summary of Thermodynamic and Kinetic Parameters for the Fyn SH3 Domain/Peptide Interaction

T (ºC)	K_D (nM)[a]	ΔH_D (kcal mol^{-1})[b]	ΔS_D (cal mol^{-1} K^{-1})[c]	k_{off} (s^{-1})[d]	k_{on} ($\times 10^8$ M^{-1} s^{-1})[e]	$D_{protein}$ ($\times 10^{-6}$ cm^2 s^{-1})[f]	$D_{peptide}$ ($\times 10^{-6}$ cm^2 s^{-1})[g]
10	37.5±0.2	11.2±0.4	5±1	4.5±0.2	1.06±0.06	1.55±0.03	0.78±0.03
20	72±5	13.5±0.3	14±2	11.7±0.8	1.4±0.2	2.1±0.1	1.08±0.02
30	158±10	19.5±0.6	33±2	27.5±0.6	1.5±0.1	2.84±0.05	1.46±0.03
35	272±19	21.8±0.6	40±2	62±2	2.0±0.2	—	—
40	442±28	23±1	43±4	117±3	2.3±0.2	—	—
45	—	—	—	175±2	—	—	—
50	1389±196	27±1	56±4	331±5	2.1±0.3	—	—

[a] Equilibrium dissociation constant determined by ITC
[b] Molar enthalpy change upon dissociation determined by ITC
[c] Molar entropy change upon dissociation determined by ITC
[d] Dissociation rate constant determined by NMR. Value at 20°C is the error-weighted average of CPMG- and ZZ-exchange-derived values
[e] Association rate constant determined by NMR and ITC according to (Eq. 6.23)
[f] Translational diffusion constant of the Fyn SH3 domain determined by NMR
[g] Translational diffusion constant of the 12 residue peptide determined by NMR

CHAPTER 5: CONCLUSION

5.1 Summary of findings

The ^{15}N and 1H chemical shifts differences between the free and bound form of the Fyn SH3 domain were measured. Residues with the largest differences are located in the vicinity of the identified binding site for canonical Type 1 ligands.

Off-rates were measured by ZZ magnetization exchange at 10ºC and 20º, and by CPMG dispersion-relaxation at 20ºC, 30ºC, 35ºC, 40ºC, 45ºC and 50ºC. Both experimental data sets could be fitted as a group for each temperature using unique values for k_{ex} and p_b without reducing the agreement between predicted and observed peak intensities, arguing that binding occurs in a two-state manner.

The proposition that peptide binding does not significantly populate intermediate states is further supported by the absence of additional peaks, the absence of dispersion-relaxation in the *apo* or saturated forms of the Fyn SH3 domain and the high correlation between chemical shift differences measured by HSQC and those extracted from CPMG experiments.

The bound population p_b was fixed over all temperatures, and chemical shift differences values were fixed from measured HSQC values so as to improve the robustness of CPMG data fittings. Rates of dissociation span two orders of magnitude, from $4.5s^{-1}$ at 10ºC to $331s^{-1}$ at 50ºC. The magnitude of these rates at higher

temperatures would make them difficult to measure with stopped flow techniques. The off-rates found at 20ºC by ZZ exchange and CPMG are in good agreement.

ITC data collected at 6 temperatures ranging from 10°C to 50°C are remarkably self-consistent and can be completely accounted for by 4 parameters: the enthalpy and entropy of dissociation, the difference in heat capacity between the free and bound states, and a correction factor for the concentration of injected peptide. The free protein and peptide have a significantly higher heat capacity than the bound complex, consistent with an increase in the exposure of non-polar groups to solvent upon dissociation.

Association rate constants follow Eyring kinetics, with a forward activation energy of (3.4±0.4) kcal/mol. The magnitudes and temperature dependence of k_{on} are consistent with a diffusion-limited process. Literature values for apparent activation enthalpies are approximately 4.5 kcal mol^{-1} for diffusive processes[66]. A similar value, (4.5±0.4) kcal mol^{-1} is found by using the experimentally-determined diffusion coefficients. Indeed, self-diffusion coefficients were measured by DOSY for the peptide and the protein and were found to follow the Stokes-Einstein relationship.

The temperature dependence of NMR-derived k_{off} values indicates that there is a significant energy barrier to dissociation. Curvature in the k_{off} Eyring plot closely matches the ITC-derived value of ΔC_p, which suggests that the hydrophobic binding interface is highly solvated upon reaching the rate-limiting-step of dissociation.

Good agreement between NMR and ITC results validates the utility of NMR as a thermodynamic technique. In addition, the comparison of the NMR and ITC results provided an element of cross-validation between the two techniques: association rate constants calculated with a combination of NMR and ITC data are physically reasonable, exhibiting both the magnitude and temperature dependence typical of a diffusion-limited reaction, while the temperature dependence of NMR-derived dissociation rates matches the ITC-derived heat capacity value. The internal consistency of the combined dataset highlights the potential of using NMR and calorimetry in concert to quantitate functional dynamics in proteins.

5.2 Biological relevance

The fast timescale binding kinetics we have measured by NMR for the SH3 domain portion of the Fyn tyrosine kinase may influence the ability of this enzyme to participate in diverse cellular functions that include T-cell activation, mitogenesis, brain function, and cell-adhesion mediated signaling[67]. Fyn belongs to the Src family of non-receptor tyrosine kinases, whose members contain an SH3 and SH2 domain separated from the catalytic domain by a linker region[68] (Fig. 1.1). In the inactive form of these enzymes, the SH3 domain binds a proline-rich target in the linker region and the SH2 domain binds a phosphotyrosine-containing sequence at the C-terminus, which locks the kinase domain in a closed conformation[77-80]. The enzyme can be activated by displacement of the intramolecular target from the SH3 domain by a proline-rich region of another protein[69,70].

In order for the activation of Fyn to be rapid, dissociation of the linker and association of another binding partner must take place on a fast timescale. Our results show that association rates for the Fyn SH3 domain can be close to the theoretical limit, matching those of cytotoxic nuclease/inhibitor complexes where rapid binding is thought to be critical for cell survival[33,71]. Fast association rates, in turn, allow dissociation rates to be large while maintaining relatively high affinity. Thus, the fast binding kinetics observed in this study are optimal for the rapid response of Fyn to changing intracellular signals.

The study of SH3 domain-peptide interactions is quite relevant in the context of drug design, where peptides are likely leads[72] and SH3 domains are likely targets[10]. Because of their prevalence in signaling pathways, SH3 domains are associated with several pathologies, among which AIDS, cancer, and various forms of leukemia.

5.3 Future directions

Characterization of the binding transition state

In this project, we have successfully constructed a model system to study the fast association of a protein and peptide. Depending on the complexity of the interaction, many local and global parameters are revealed by the use of NMR and ITC experiments. The combined dataset also provides detailed information not accessible from either the NMR or ITC measurements alone. For instance, with the combined NMR and ITC dataset, we calculated association rate constants, and from their temperature dependence characterized the rate-determining step of association.

The techniques that were applied in our study are readily portable to other bimolecular systems, as long as one of the components presents favourable NMR qualities and that

the interaction and experimental conditions are amenable to ITC. A natural extension of our study is to slightly change the experimental conditions so as to probe various aspects of the binding mechanism.

For instance, electrostatic interactions can be strengthened or weakened by changing the pH and ionic strength of the sample buffer, altering the electrostatic enhancement factor of association. The isolated contribution of Site 1 to electrostatic enhancement (Fig.. 1.4), in the case of SH3 domains, can be probed by competing away the peptide with free arginine. Similarly, addition of denaturant assesses the importance of hydrophobic desolvation.

The diffusion-dependence of the interaction[73] can be assessed by addition of a viscogen to the solution; this would allow one to independently test for diffusive and energetic barriers, rather than changing the viscosity indirectly by varying the temperature. It has been reported that molecular crowding can in some circumstances enhance association rates through size-exclusion effect[74]. It would be interesting to do a systematic survey of those parameters using a combination of viscogens and crowding agents of various sizes.

For each modification of the experimental conditions, it is important to evaluate the changes that can affect the free components, the transition state or the bound complex. For instance, if the ionic strength is to be modified, salts can be added incrementally into the free Fyn protein and structural changes can be monitor by recording a series of ^{15}N-edited HSQC. A similar procedure has to be employed on the peptide-saturated Fyn sample.

Finally, modification of the ligand and protein sequences could provide a tool similar to Φ-value analyses, revealing the extent to which interactions are formed in the TS for energy-barrier limited systems. The peptide ligand could incorporate unnatural amino acids or chemical modifications to test hypotheses about the nature of the binding mechanism, for instance to assess the role of conformational entropy.

Multiple binding motifs

Not all PxxP motifs occur in isolation: a search in the human transcriptome reveals that >400 transcript variants contain multiple canonical SH3 motifs. For instance, the tau protein, normally responsible for association with and stabilization of microtubules, possesses two PxxP SH3 recognition motifs in close proximity: class I motif PTPPTR and class II-like motif RTPPKSP. Both sequences are essential for high affinity binding as demonstrated using truncation mutants[75]. The biophysical significance of tandem PxxP sequences is not well understood, despite their abundance in expressed proteins.

To tackle this question, the Fyn/tau interaction could be studied by characterizing the binding of the PxxP sequences, both isolated and in tandem. Ultimately, interactions of the Fyn SH3 domain with full-length tau could be studied if the spectral quality is not altered by binding. The experiments would yield thermodynamic, kinetic, and structural information on the interaction and will reveal whether the PxxP motifs bind to the Fyn SH3 domain simultaneously or compete for the same binding site. It could also be determined how repetition of the PxxP sequences affects the affinity, specificity and binding kinetics of the interaction. These results will have profound implications for the design of inhibitors. It has to be noted that the interaction between the Fyn SH3 domain and Tau protein has been closely associated with Alzheimer's disease.

APPENDIX: Materials and Methods

6.1 Peptide ligands

A peptide with the sequence Ac-VSLARRPLPPLP-NH_2 was synthesized by the Sheldon Biotechnology Centre (McGill, Montreal), and purified to >98% homogeneity by reverse-phase HPLC. A MW of 1356.85 Da was confirmed by mass spectrometry. The peptide powder was dissolved in water and dialyzed against 20 mM sodium phosphate, 1 mM EDTA, pH 6.0. The concentration of the dialyzed solution was estimated by amino acid analysis (The Hospital for Sick Children, Toronto). Samples were hydrolyzed in 6M HCl, 1% phenol for 24 h at 110ºC, derivatized with phenyisothiocyanate to produce phenylthiocarbamyl amino acids, and quantified by reverse-phase HPLC.

6.2 Protein Sample Preparation

The protein samples used in this study were produced with a pET11d-based expression plasmid (Novagen) coding for the SH3 domain (residues 81 to 148) of the chicken isoform of the Fyn tyrosine kinase fused to an N-terminal 6-histidine affinity purification tag. Proteins for ITC studies were expressed in freshly-transformed *E. coli* BL21(DE3) bacteria grown at 37ºC in Luria broth (LB). Expression was induced at an OD_{600} of 0.8 by adding 200 mg L^{-1} IPTG (isopropyl β-D-1-thiogalactopyranoside). Cells were harvested by centrifugation 3 hours later. Isotopically-labeled proteins were expressed according to the protocol of Marley *et al*[76]. Bacteria were first cultured in 4 L of LB to an OD_{600} of 0.8, then transferred to 1 L of M9 minimal media, which contained $^{15}NH_4Cl$ and D-glucose as the sole sources of nitrogen and carbon. In the case of samples used for triple-resonance assignment experiments, the D-glucose was uniformly ^{13}C labeled.

Expression was induced 1 hour after the transfer to M9 media and cells were harvested by centrifugation after a further 3 hours. Protein purification was performed with a modified protocol based on that of Maxwell and Davidson[8]. Cell pellets were resuspended in solution containing 0.1 M NaH_2PO_4, 0.01 M Tris, 6 M Gdm-Hcl, and 0.01 M imidazole, at pH 8.0 and rocked for 1 hour at room temperature to achieve lysis. The lysate was centrifuged, and the supernatant was applied to a Toyopearl AF-Chelate-650M column (Tosoh Bioscience). The column was washed extensively with lysis buffer and the protein was eluted with a solution containing 0.2 M acetic acid and 6 M Gdm-HCl. The eluate was dialyzed against a solution containing 150 mM NaCl, 50 mM Tris, 2 mM EDTA, and 5 mM benzamidine, at pH 7.9, and purified by size-exclusion chromatography using a Superdex 75 10/300 GL column (GE HealthSciences). The elution volume of the protein corresponded to the nominal molecular weight of the desired product, 8.7 kDa. Sample purity was estimated by SDS-PAGE to be greater than 99%. Protein concentration was estimated using a theoretical 280 nm extinction coefficient[77].

6.3 NMR spectroscopy

All NMR experiments were performed on protein samples containing 20 mM sodium phosphate, pH 6.0, 1mM EDTA, 10 μM DSS (dimethyl-silapentane-sulfonate) 0.02% sodium azide, and 10% D_2O. NMR spectra were recorded on Varian INOVA spectrometers operating at 800 MHz and 500 MHz proton Larmor frequencies, (18.8 T and 11.7 T), equipped with high-sensitivity cold-probes. All 2D spectra were processed with shifted sine-bell weighting functions in the t_1 (^{15}N) and t_2 (^{1}H) dimensions. Chemical shifts and peak intensities were fitted using the nlinLS program of the nmrPipe package[78].

6.4 Spectral Assignment

All 66 backbone amide resonances, except those of the 6-histidine tag, were assigned using HNCACB spectra[79], recorded at 30ºC and 18.8 T for the Fyn SH3 domain in both the peptide-free (0.34 mM Fyn) and peptide-saturated (0.34 mM Fyn, 0.5 mM peptide) states. $^{1}H-^{15}N$ HSQC spectra[80] of the free and peptide-saturated Fyn SH3 domain were recorded at 10ºC, 20ºC, 30°C, 35ºC, 40ºC, 45°C and 50ºC. ^{1}H and ^{15}N chemical shifts varied approximately linearly with temperature such that the assignments obtained at 30ºC could be propagated to the other temperatures. Accordingly, the ^{15}N chemical shift difference, $\Delta\delta_{HSQC}$, between the free and peptide-saturated resonances of the protein could be computed for all temperatures.

6.5 Monte Carlo Error Analysis

Monte Carlo (MC) error analyses were performed by repeating minimization calculations on MC datasets generated according to

$$X_j^{MC} = X_j^{\exp} + \sigma_j \varepsilon \qquad \text{(Eq. 6.1)}$$

where X_j is the jth experimental datum with experimental uncertainty σ_j and ε is drawn randomly from a normal distribution with a mean of zero and standard deviation of one. The standard deviations of parameters extracted from the MC fits were taken as the experimental uncertainties of the parameters. Errors in ΔH_0, ΔS_0, ΔC_p, $\Delta H_A^{\ddagger}$, $\Delta H^{\ddagger}_{D,}$ r_h, and ZZ-exchange-derived kinetic parameters were obtained with this method.

6.6 ZZ-exchange Experiments

ZZ exchange NMR experiments[40] were performed at 10ºC and 20ºC at a proton Larmor frequency of 800 MHz, with 64 and 833 complex points in the indirect and direct dimensions, collecting 32 scans for each indirect point. Mixing times, T_{mix} (labeled T in Farrow *et al*) of (0, 0.0214, 0.0482, 0.0856, 0.1445, 0.1927, 0.3425, 0.5352, 0.7707, 1.0490) seconds and (0, 0.0214, 0.0482, 0.0856, 0.1445, 0.1927, 0.3425, 0.5352, 0.7707) seconds were used at 10°C and 20°C, respectively. Data were fit numerically to an equation that accounts for exchange during the magnetization transfer steps:

$$\begin{bmatrix} I_{ff} & I_{bf} \\ I_{fb} & I_{bb} \end{bmatrix} = \begin{bmatrix} 1 & y_{bf} \\ y_{fb} & y_{bb} \end{bmatrix} \bullet \exp\left(T_{mix} \begin{bmatrix} -k'_{on} - R_1^f & k_{off} \\ k'_{on} & -k_{off} - R_1^b \end{bmatrix} \right) \bullet \begin{bmatrix} I_0^f & 0 \\ 0 & I_0^b \end{bmatrix} \quad \text{(Eq. 6.2)}$$

$$k'_{on} = k_{on}[X] = k_{off} \frac{p_B}{1 - p_B} \quad \text{(Eq. 6.3)}$$

where I_0^f and I_0^b are the amplitudes of the free and bound signals immediately before the mixing delay, R_1^f and R_1^b are the ^{15}N longitudinal relaxation rates in the free and bound states, y_{bf} and y_{fb} account for exchange and y_{bb} accounts for differential relaxation during the magnetization transfer steps that follow the mixing delay. I_{ff} and I_{bb} are the experimentally-observed intensities of the free and bound auto-peaks, and I_{fb} and I_{bf} are the experimentally-observed intensities of the cross-peaks. The matrix exponential was calculated numerically in MATLAB. The parameters I_0^f, I_0^b, y_{fb}, y_{bf}, y_{bb}, R_1^f, and R_1^b were optimized on a per-residue basis. The values of k_{off} and p_B were optimized twice, once on a per-residue basis, and again as global parameters for all residues at a given temperature. Non-linear fits were performed by minimizing the residual sum of squares

(RSS) between experimental peak intensities and those back-calculated using (Eq. 6.2). The RSS for an individual residue is given by

$$RSS = \sum_{T_{mix}} \sum_{j=ff,bb,fb,bf} \left[I_j^{calc}(T_{mix}) - I_j^{exp}(T_{mix}) \right]^2 \qquad \text{(Eq. 6.4)}$$

The experimental uncertainties of k_{off} and p_B were estimated using Monte Carlo (MC) simulations with 830 iterations, according to (Eq. 6.1) and (Eq. 6.2) with global exchange parameters. Uncertainties in peak intensities, σ_I, were estimated on a per-residue basis using the RSS from fits with individually-optimized values k_{off} and p_B,

$$\sigma_I = \sqrt{\frac{RSS}{DF}} \qquad \text{(Eq. 6.5)}$$

where DF is the number of degrees of freedom of the per-residue fit, 31 for data at 10°C and 27 at 20°C.

6.7 CPMG Measurements

^{15}N CPMG single-quantum relaxation dispersion experiments were performed at 20°C, 30°C, 35°C, 40°C, 45°C, and 50°C using a recently-published pulse sequence that utilizes proton decoupling during the CPMG period[41]. Data were collected at 500 (800) MHz proton Larmor frequencies with 64 (80) and 512 (833) complex points in the indirect and direct dimensions, respectively, averaging 16 scans for each complex indirect point, with a constant relaxation delay, T_{relax}, of 40 ms. Between 12 and 24 unique values of ν_{CPMG} covering a range of 25 Hz to 1000 Hz were used at each temperature and static magnetic field strength, where $\nu_{CPMG}=1/(2\tau_{CP})$, and τ_{CP} is the delay between successive refocusing pulses in the CPMG pulse train. The transverse relaxation rate, R_2, was calculated for each ν_{CPMG} value according to

$$R_2\left(\nu_{CPMG}\right) = -\frac{1}{T_{relax}} \ln\left\{\frac{I\left(\nu_{CPMG}\right)}{I_0}\right\} \qquad \text{(Eq. 6.6)}$$

where $I(\nu_{CPMG})$ is the peak intensity obtained with a given CPMG pulse spacing and I_0 is the intensity of the peak when the relaxation delay is omitted, $T_{relax}=0$. The uncertainty in peak intensity, σ_I, was assumed to be constant across a single dispersion profile, i.e. the same at all ν_{CPMG} values for a single residue, temperature, and static magnetic field strength. Values of σ_I were calculated for each dispersion profile using replicate spectra collected at several different ν_{CPMG} values, according to the definition of the pooled standard deviation[81]. If replicate experiments were performed at N_{dup} different ν_{CPMG} values, such that n_j replicates were obtained at $(\nu_{CPMG})_j$ and the intensities of a given peak in these n_j spectra have a standard deviation s_j, then the pooled standard deviation in peak intensity, σ_I, was calculated as

$$\sigma_I^2 = \frac{\sum_{j=1}^{N_{dup}} s_j^2\left(n_j - 1\right)}{\sum_{j=1}^{N_{dup}} \left(n_j - 1\right)} \qquad \text{(Eq. 6.7)}$$

The corresponding errors in transverse relaxation rates, σ_{R2}, across the profile are thus

$$\sigma_{R2}\left(\nu_{CPMG}\right) = \frac{\sigma_I I_0}{T_{relax} I\left(\nu_{CPMG}\right)} \qquad \text{(Eq. 6.8)}$$

CPMG data were fit using software generously provided by Dr.s Korzhnev and Kay, which generates dispersion profiles numerically according to[47],

$$R_2^{calc} = -\frac{1}{T_{relax}} \ln\left\{\frac{M_B\left(2n_{CP}\tau_{CP}\right)}{M_B\left(0\right)}\right\} \qquad \text{(Eq. 6.9)}$$

where $2n_{CP}$ refocusing pulses separated by delays of τ_{CP} are applied during the relaxation period, $T_{relax} = 2n_{CP}\tau_{CP}$. M_F and M_B are the signals arising from the bound and free states, and are calculated following the modified Bloch-McConnell equations,

$$\begin{bmatrix} M_F(2n_{CP}\tau_{CP}) \\ M_B(2n_{CP}\tau_{CP}) \end{bmatrix} = \left(\exp\left(\mathbf{A}\tfrac{1}{2}\tau_{CP}\right)\exp\left(\tilde{\mathbf{A}}\tau_{CP}\right)\exp\left(\mathbf{A}\tfrac{1}{2}\tau_{CP}\right)\right)^{n_{CP}} \begin{bmatrix} M_F(0) \\ M_B(0) \end{bmatrix} \quad \text{(Eq. 6.10)}$$

where

$$\mathbf{A} = \begin{bmatrix} -k'_{on} - R_2^{\infty} + i\Delta\omega & k_{off} \\ k'_{on} & -k_{off} - R_2^{\infty} \end{bmatrix} \quad \text{(Eq. 6.11)}$$

$$\tilde{\mathbf{A}} = \begin{bmatrix} -k'_{on} - R_2^{\infty} - i\Delta\omega & k_{off} \\ k'_{on} & -k_{off} - R_2^{\infty} \end{bmatrix} \quad \text{(Eq. 6.12)}$$

k_{off} is the dissociation rate constant, k'_{on} is defined in (Eq. 6.3), $\Delta\omega$ is the difference in precession frequency between the free and bound states, expressed in radians s^{-1}, and R_2^{∞} is the transverse relaxation rate in the absence of conformational exchange, which was assumed to be equal for the free and bound states. The parameters k_{off}, p_B, $\Delta\omega$, and R_2^{∞} were fit by minimizing the chi-squared function

$$\chi^2 = \sum_{\nu_{CPMG}} \frac{\left(R_2^{exp}(\nu_{CPMG}) - R_2^{calc}(\nu_{CPMG})\right)^2}{\sigma_{R2}(\nu_{CPMG})^2} \quad \text{(Eq. 6.13)}$$

according to Models 1, 2 and 3, as described in the text. Errors in extracted parameters were estimated from the inverted Hessian matrix. In Models 1 and 2, under conditions of fast exchange, values of k_{ex} and Φ_{ex} were calculated from the extracted values of k_{off}, p_B, and $\Delta\omega$ according to (Eq. 2.2) and (Eq. 2.4). A residue was selected for analysis at a

given temperature if its data at 18.8 T and 11.7 T met the following criteria: i) the magnitude of the dispersion, $\max\left(R_2^{exp}\right) - \min\left(R_2^{exp}\right)$ is greater than 1.0 s^{-1}; ii) the value of χ^2 obtained with Model A is less than 4-fold larger than the number of degrees of freedom (# data points - # fitted parameters); iii) The fit obtained with (Eq. 6.10) is better than the fit provided by a horizontal line, at a 99.9% confidence limit. Following these criteria, 19, 32, 25, 20, 20, and 22 residues were used in the analysis at 20, 30, 35, 40, 45, and 50°C, respectively. In Model 3, optimization of the bound state population was accomplished by repeating fits with $\Delta\omega_{CPMG}$ set according to $\Delta\delta_{HSQC}$ values, and p_B fixed at a 0.01 intervals between 0.66 and 0.86. The the sum of χ^2 values over all temperatures shows a parabolic dependence on p_B with a minimum at p_B=0.74.

6.8 Diffusion measurements

Translational diffusion constants were measured with a room-temperature probe at 10°C, 20°C, and 30°C and a proton Larmor frequency of 500 MHz, using a 1D ^{1}H linear-encode-decode (LED) pulsed-field-gradient (PFG) experiment[53] and samples containing either the protein alone or peptide alone. The diffusion constant was extracted from a regression plot of $\ln\{I\}$ versus G_z^2. Integration of peak intensities was done in VNMR using a spline baseline correction.

6.9 Isothermal Titration Calorimetry

Titrations of peptide into a protein solution were carried at 10ºC, 20ºC, 30ºC, 35ºC, 40ºC and 50ºC using a VP-ITC instrument (MicroCal LLC, Northampton, MA). Both peptide and protein solutions contained 20 mM sodium phosphate at pH 6.0 and 1mM EDTA. At 20ºC to 50ºC, the concentrations of peptide and protein were approximately

0.2 mM and 0.02 mM, respectively. Experiments were done in triplicate at each temperature and consisted of 56 injections of 5 μL. At 10ºC, the concentrations of peptide and protein were approximately 0.04 mM and 0.004 mM, respectively. Experiments were done in duplicate, and consisted of 28 injections of 10 μL. Injections at all temperatures were made at a rate of 0.5 μL s^{-1} and at intervals of 300 s. The first injection peak was discarded from the isotherm, as were injection peaks without a stable baseline. The baseline was generated automatically by the MicroCal Origin package and corrected manually. Offsets were adjusted so that the isotherms approached 0 kcal mol^{-1} at saturating peptide concentrations. Isotherms were fitted using the One Set of Sites model in the MicroCal Origin package, which yielded the effective number of binding sites per protein (N_b), equilibrium dissociation constant (K_D), enthalpy of dissociation (ΔH_D) together with standard errors of the fitted parameters. The entropy of dissociation, ΔS_D, is calculated as

$$\Delta S_D = \frac{\Delta H_D}{T} + R \ln K_D \qquad \text{(Eq. 6.14)}$$

and

$$\sigma_S = \sqrt{\left(\frac{\sigma_H}{T}\right)^2 + \left(\frac{R\sigma_K}{K_D}\right)^2} \qquad \text{(Eq. 6.15)}$$

where σ_S, σ_H and σ_K are the standard errors of the dissociation entropy, enthalpy and equilibrium dissociation constant, respectively. The values of ΔH_D, ΔS_D and K_D used in the subsequent analysis were calculated as the error-weighted mean values of replicate measurements, and the experimental uncertainties were computed accordingly. For a set

of N measurements, X_1 to X_N, with standard errors, $\sigma_{X,1}$ to $\sigma_{X,N}$, the error-weighted mean is given by

$$\bar{X} = \frac{\sum_{j=1}^{N} X_j \left(\sigma_{X,j}\right)^{-2}}{\sum_{j=1}^{N} \left(\sigma_{X,j}\right)^{-2}} \qquad \text{(Eq. 6.16)}$$

and the standard error of the mean is

$$\sigma_{\bar{X}} = \frac{1}{\sqrt{\sum_{j=1}^{N} \left(\sigma_{X,j}\right)^{-2}}} \qquad \text{(Eq. 6.17)}$$

In order to extract accurate values of ΔH_0, ΔS_0, and ΔC_p from the ITC-derived parameters, a correction factor for the concentration of injected peptide was included as an adjustable parameter in the minimizations. The fitted values of ΔH_D and K_D depend on the concentration of injected peptide, such that if the estimated peptide concentration differs from the true value by a factor of β, then the fitted values of ΔH_D and K_D differ from the true values by factors of β^{-1} and β, respectively. This relationship emerges from the expression for the total enthalpy of the protein throughout an ITC titration:

$$H\left([X]_t\right) - H(0) = -\frac{\Delta H_D V}{2}\left(N_b[P]_t + [X]_t + K_D - \sqrt{\left(N_b[P]_t + [X]_t + K_D\right)^2 - 4[X]_t N_b[P]_t} \right) \qquad \text{(Eq. 6.18)}$$

where $[X]_t$ and $[P]_t$ are the total concentrations of peptide and protein, respectively, at any point in the titration, $H(0)$ is the initial total protein enthalpy, $H([X]_t)$ is the total protein enthalpy at a given peptide concentration, N_b is the number of binding sites per protein, and V is the volume of the sample cell. The enthalpy profile corresponding to the

parameters $[X]_t$, K_D, ΔH_D, and N_b is superimposable on one corresponding to $[X]_t' = \beta[X]_t$, $K_D' = \beta K_D$, $\Delta H_D' = \beta^{-1}\Delta H_D$, and $N_b' = \beta N_b$. We have verified that this equivalence also holds for calculated ITC isotherms, when sample dilution and the fixed volume of the sample cell are taken into account. The relevant expressions are significantly longer than (Eq. 6.18), and are not shown here. Because all ITC experiments were performed with the same stock of dissolved peptide, it was only necessary to fit a single peptide concentration correction factor, β, for all temperatures. The values of ΔH_0, ΔS_0, ΔC_p, and β were extracted by minimizing the chi-squared function

$$\chi^2 = \sum_T \frac{\left(\Delta H_D^{calc}(T) - \Delta H_D^{\exp}(T)\right)}{\sigma_H^2(T)} + \sum_T \frac{\left(\Delta S_D^{calc}(T) - \Delta S_D^{\exp}(T)\right)}{\sigma_S^2(T)} + \sum_T \frac{\left(K_D^{calc}(T) - K_D^{\exp}(T)\right)}{\sigma_K^2(T)},$$

(Eq. 6.19)

$$\Delta H_D^{calc}(T) = \beta^{-1}\left(\Delta H_0 + \Delta C_p\left(T - T_0\right)\right) \qquad \text{(Eq. 6.20)}$$

$$K_D^{calc}(T) = \beta \exp\left\{\frac{-1}{RT}\left(\Delta H_0 + \Delta C_p\left(T - T_0\right) - T\left(\Delta S_0 + \Delta C_p \ln\left\{\frac{T}{T_0}\right\}\right)\right)\right\} \qquad \text{(Eq. 6.21)}$$

$$\Delta S_D^{calc}(T) = \frac{\Delta H_D^{calc}(T)}{T} + R \ln K_D^{calc}(T) \qquad \text{(Eq. 6.22)}$$

Errors in thermodynamic parameters were estimated by Monte Carlo simulations with 2840 iterations, using (Eq. 6.1), (Eq. 6.20), (Eq. 6.21), and (Eq. 6.22). The extracted value of β=0.88 indicates that the peptide concentration determined by amino acid analysis underestimated the true concentration by approximately 12%. The χ^2 obtained with β=1, i.e. no peptide concentration adjustment, is about twice that with β=0.88. Note that experimental, i.e. uncorrected, thermodynamic parameters are listed in Table 4.1 and plotted in Figure 3.2. The true values of ΔH_D and K_D are related to the experimental values by factors of β and β^{-1}, respectively. A corrected value of the equilibrium

dissociation constant, K_D^{corr}, was used to calculate the association rate constant according to

$$k_{on} = \frac{k_{off}}{K_D^{corr}} = \frac{k_{off}}{\beta^{-1} K_D} \qquad \text{(Eq. 6.23)}$$

where K_D is the experimental equilibrium dissociation constant.

BIBLIOGRAPHY

(1) Palmer, A. G.; Kroenke, C. D.; Loria, J. P. *Methods in Enzymology* **2001**, *339*, 204-238.

(2) Preisinger, C.; von Kriegsheim, A.; Matallanas, D.; Kolch, W. *Proteomics* **2008**.

(3) Gabdoulline, R. R.; Wade, R. C. *Current Opinion in Structural Biology* **2002**, *12*, 204-213.

(4) Schreiber, G.; Fersht, A. R. *Biochemistry* **1993**, *32*, 5145-5150.

(5) Laidler, K. J.; Meiser, J. H.; Sanctuary, B. C. *Physical chemistry*; Houghton Mifflin: Boston :, 2003.

(6) Park, H. Y.; Qiu, X.; Rhoades, E.; Korlach, J.; Kwok, L. W.; Zipfel, W. R.; Webb, W. W.; Pollack, L. *Anal Chem* **2006**, *78*, 4465-73.

(7) Mittermaier, A.; Kay, L. E. *Science* **2006**, *312*, 224-8.

(8) Maxwell, K. L.; Davidson, A. R. *Biochemistry* **1998**, *37*, 16172-16182.

(9) Larson, S. M.; Davidson, A. R. *Protein Sci* **2000**, *9*, 2170-80.

(10) Dalgarno, D. C.; Botfield, M. C.; Rickles, R. J. *Biopolymers* **1997**, *43*, 383-400.

(11) Mayer, B. J. *Journal of Cell Science* **2001**, *114*, 1253-1263.

(12) Noble, M. E.; Musacchio, A.; Saraste, M.; Courtneidge, S. A.; Wierenga, R. K. *Embo J* **1993**, *12*, 2617-24.

(13) Morton, C. J.; Pugh, D. J.; Brown, E. L.; Kahmann, J. D.; Renzoni, D. A.; Campbell, I. D. *Structure* **1996**, *4*, 705-14.

(14) Renzoni, D. A.; Pugh, D. J. R.; Siligardi, G.; Das, P.; Morton, C. J.; Rossi, C.; Waterfield, M. D.; Campbell, I. D.; Ladbury, J. E. *Biochemistry* **1996**, *35*, 15646-15653.

(15) Arold, S.; O'Brien, R.; Franken, P.; Strub, M. P.; Hoh, F.; Dumas, C.; Ladbury, J. E. *Biochemistry* **1998**, *37*, 14683-91.

(16) Fernandez-Ballester, G.; Blanes-Mira, C.; Serrano, L. *J Mol Biol* **2004**, *335*, 619-29.

(17) Rickles, R. J.; Botfield, M. C.; Zhou, X. M.; Henry, P. A.; Brugge, J. S.; Zoller, M. J. *Proc. Natl Acad. Sci.* **1995**, *92*, 10909–10913.

(18) Li, H.; Lawrence, D. S. *Chem Biol* **2005**, *12*, 905-12.

(19) Sandström, J. *Dynamic NMR spectroscopy*; Academic Press: London ; New York :, 1982.

(20) Sugase, K.; Dyson, H. J.; Wright, P. E. *Nature* **2007**, *447*, 1021-U11.

(21) Berg, O. G.; Vonhippel, P. H. *Annual Review of Biophysics and Biophysical Chemistry* **1985**, *14*, 131-160.

(22) Alsallaq, R.; Zhou, H. X. *Proteins-Structure Function and Bioinformatics* **2008**, *71*, 320-335.

(23) Selzer, T.; Schreiber, G. *Journal of Molecular Biology* **1999**, *287*, 409-419.

(24) Vijayakumar, M.; Wong, K. Y.; Schreiber, G.; Fersht, A. R.; Szabo, A.; Zhou, H. X. *Journal of Molecular Biology* **1998**, *278*, 1015-1024.

(25) Berg, O. G. *Biophysical Journal* **1985**, *47*, 1-14.

(26) Solc, K.; Stockmayer, W. *Int. J. Chem. Kin.* **1973**, *5*, 733-752.

(27) Gabdoulline, R. R.; Wade, R. C. *Journal of Molecular Recognition* **1999**, *12*, 226-234.

(28) Di Nardo, A. A.; Larson, S. M.; Davidson, A. R. *Journal of Molecular Biology* **2003**, *333*, 641-655.

(29) Chan, B.; Lanyi, A.; Song, H. K.; Griesbach, J.; Simarro-Grande, M.; Poy, F.; Howie, D.; Sumegi, J.; Terhorst, C.; Eck, M. J. *Nat Cell Biol* **2003**, *5*, 155-60.

(30) Viguera, A. R.; Arrondo, J. L.; Musacchio, A.; Saraste, M.; Serrano, L. *Biochemistry* **1994**, *33*, 10925-33.

(31) Zamora-Leon, S. P.; Lee, G.; Davies, P.; Shafit-Zagardo, B. *J Biol Chem* **2001**, *276*, 39950-8.

(32) Sydor, J. R.; Engelhard, M.; Wittinghofer, A.; Goody, R. S.; Herrmann, C. *Biochemistry* **1998**, *37*, 14292-14299.

(33) Wallis, R.; Moore, G. R.; James, R.; Kleanthous, C. *Biochemistry* **1995**, *34*, 13743-13750.

(34) Northrup, S. H.; Thomasson, K. A.; Miller, C. M.; Barker, P. D.; Eltis, L. D.; Guillemette, J. G.; Inglis, S. C.; Mauk, A. G. *Biochemistry* **1993**, *32*, 6613-6623.

(35) England, P.; Bregegere, F.; Bedouelle, H. *Biochemistry* **1997**, *36*, 164-172.

(36) Altschuh, D.; Dubs, M. C.; Weiss, E.; Zederlutz, G.; Vanregenmortel, M. H. V. *Biochemistry* **1992**, *31*, 6298-6304.

(37) Stone, S. R.; Dennis, S.; Hofsteenge, J. *Biochemistry* **1989**, *28*, 6857-6863.

(38) Taylor, M. G.; Rajpal, A.; Kirsch, J. F. *Protein Science* **1998**, *7*, 1857-1867.

(39) Radic, Z.; Kirchhoff, P. D.; Quinn, D. M.; McCammon, J. A.; Taylor, P. *Journal of Biological Chemistry* **1997**, *272*, 23265-23277.

(40) Farrow, N. A.; Zhang, O.; Forman-Kay, J. D.; Kay, L. E. *J. Biomol. NMR* **1994**, *4*, 727-734.

(41) Hansen, A. F.; Vallurupalli, P.; Kay, L. E. *Journal of Physical Chemistry B* **2008**, *112*, 5898-5904.

(42) Luz, Z.; Meiboom, S. *J. Chem. Phys.* **1963**, *39*, 366.

(43) Carver, J. P.; Richards, R. E. *J. Magn. Reson.* **1972**, *6*, 89-105.

(44) Tollinger, M.; Skrynnikov, N. R.; Mulder, F. A.; Forman-Kay, J. D.; Kay, L. E. *J. Am. Chem. Soc* **2001**, *123*, 11341-11352.

(45) Mittag, T.; Schaffhausen, B.; Gunther, U. L. *Journal of the American Chemical Society* **2004**, *126*, 9017-9023.

(46) Mittag, T.; Schaffhausen, B.; Gunther, U. L. *Biochemistry* **2003**, *42*, 11128-11136.

(47) Korzhnev, D. M.; Salvatella, X.; Vendruscolo, M.; Di Nardo, A. A.; Davidson, A. R.; Dobson, C. M.; Kay, L. E. *Nature* **2004**, *430*, 586-590.

(48) Tolkatchev, D.; Xu, P.; Ni, F. *Journal of the American Chemical Society* **2003**, *125*, 12432-12442.

(49) Ferreon, J. C.; Hilser, V. J. *Biochemistry* **2004**, *43*, 7787-7797.

(50) Palencia, A.; Cobos, E. S.; Mateo, P. L.; Martinez, J. C.; Luque, I. *Journal of Molecular Biology* **2004**, *336*, 527–537.

(51) Casares, S.; AB, E.; Eshuis, H.; Lopez-Mayorga, O.; van Nuland, N. A. J.; Conejero-Lara, F. *BMC Structural Biology* **2007**, *7*, 22-40.

(52) Prabhu, N.; Sharp, K. *Annu. Rev. Phys. Chem.* **2005**, *56*, 521-548.

(53) Altieri, A. S.; Hinton, D. P.; Byrd, R. A. *Journal of the American Chemical Society* **1995**, *117*, 7566-7567.

(54) Wilkins, D. K.; Grimshaw, S. B.; Receveur, V.; Dobson, C. M.; Jones, J. A.; Smith, L. J. *Biochemistry* **1999**, *38*, 16424-16431.

(55) Castellan, G. *Physical Chemistry*; 3rd ed.; Benjamin Cummings: Menlo Park, California, 1983.

(56) Pilling, M.; Seakins, P. *Reaction Kinetics*; Oxford University Press: Oxford, 1995.

(57) Hansen, D. F.; Vallurupalli, P.; Lundstrom, P.; Neudecker, P.; Kay, L. E. *J. Am. Chem. Soc.* **2008**, *130*, 2667-2675.

(58) Vallurupalli, P.; Hansen, D. F.; Stollar, E.; Meirovitch, e.; Kay, L. E. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 18473–18477.

(59) Bhaskar, K.; Yen, S. H.; Lee, G. *Journal of Biological Chemistry* **2005**, *280*, 35119-35125.

(60) Kim, H. I.; Jung, J.; Lee, E. S.; Kim, Y. C.; Lee, W.; Lee, S. T. *Biochemical and Biophysical Research Communications* **2007**, *362*, 829-834.

(61) Shelton, H.; Harris, M. *Virology Journal* **2008**, *5*.

(62) Solomaha, E.; Szeto, F. L.; Yousef, M. A.; Palfrey, C. *Journal of Biological Chemistry* **2005**, *280*, 23147-23156.

(63) Vidal, M.; Goudreau, N.; Cornille, F.; Cussac, D.; Gincel, E.; Garbay, C. *Journal of Molecular Biology* **1999**, *290*, 717-730.

(64) Schuck, P. *Biophysical Journal* **1996**, *70*, 1230-1249.

(65) Schuck, P.; Minton, A. P. *Analytical Biochemistry* **1996**, *240*, 262-272.

(66) Gabdoulline, R. R.; Wade, R. C. *Methods* **1998**, *14*, 329-341.

(67) Resh, M. D. *Int. J. Biochem. Cell. Bio.* **1998**, *30*, 1159-1162.

(68) Parsons, S. J.; Parsons, J. T. *Oncogene* **2004**, *23* 7906-7909.

(69) Moarefi, I.; LaFevre-Bernt, M.; Sicheri, F.; Huse, M.; Lee, C. H.; Kuriyan, J.; Miller, W. T. *Nature* **1997**, *385*, 650-653.

(70) Lerner, E. C.; Smithgall, T. E. *Nat. Struct. Biol.* **2002**, *9*, 365-369.

(71) Schreiber, G.; Fersht, A. R. *Nature Structural Biology* **1996**, *3*, 427-431.

(72) Drahl, C. *Chemical & Engineering News* **2008**, *86*, 18-23.

(73) Frisch, C.; Fersht, A.; Schreiber, G. *J. Mol. Biol.* **2001**, *308*, 69-77.

(74) Kozer, N.; Kuttner, Y. Y.; Haran, G.; Schreiber, G. *Biophys J* **2007**, *92*, 2139-49.

(75) Lee, G.; Newman, S. T.; Gard, D. L.; Band, H.; Panchamoorthy, G. *J Cell Sci* **1998**, *111 (Pt 21)*, 3167-77.

(76) Marley, J.; Lu, M.; Bracken, C. *Journal of Biomolecular Nmr* **2001**, *20*, 71-75.

(77) Pace, C. N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. *Protein Science* **1995**, *4*, 2411-2423.

(78) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277-293.

(79) Wittekind, M.; Mueller, L. *J. Magn. Reson. Series B* **1993**, *101*, 201-205.

(80) Kay, L. E.; Keifer, P.; Saarinen, T. *J. Am. Chem. Soc.* **1992**, *114*, 10663-10665.

(81) McNaught, A. D.; Wilkinson, A.; 2 ed.; Royal Society of Chemistry: Cambridge, UK, 1997.