Towards the Full Molecular Investigation of Protein Kinase A Mediated Catalysis by NMR Spectroscopy

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Larry Raou Masterson

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Dr. Gianluigi Veglia

March 2009

Acknowledgements

I have been blessed by the company of many passionate scientists with whom I have had the opportunity to collaborate with. I am first thankful for the guidance and support of my advisor Gianluigi Veglia who has been a dedicated teacher and friend. I am also thankful for the initial guidance of my undergraduate research advisors, Adrienne Loh and Ronald McKelvey, who first ignited my interest in solution NMR spectroscopy of proteins. I would like to thank Professor David Thomas for his insightful discussions at Membrane Journal Club, a platform that has indeed pushed me to learn much about the many techniques used in science. I am also deeply indebted to the patience, guidance, and productive collaboration cultivated during my thesis work with Dr. Marco Tonelli at NMRFAM. I owe much to Professor George Barany who gave full support to my acceptance into the graduate program at the University of Minnesota's Department of Chemistry, although I had never even taken any of their required GRE examinations! Several others have contributed in many ways to the work described in this dissertation; they are Alessandro Mascioni, Fernando Porcelli, Dan Mullen, Michael Mueller, Warren Gallagher, and Oanh Nguyen.

And not the least of any of these acknowledgements: numerous insightful discussions and memorable moments were shared with my colleagues in the Veglia laboratory. Because of this, they have become my close friends: Nathaniel Traaseth, Kim Ha, Raffaello Verardi, Martin Gustavsson, Tao Yu, and Lei Shi.

Dedication

Countless late nights, early mornings, and long weekends were spent along this journey. Therefore, I dedicate my doctoral work to my wife who stood by my side all along the way. This work would not have been possible without the love and endurance displayed during this time.

Abstract

The reversible phosphorylation of proteins is fundamental to the modulation of myocardial contraction. A mechanism which controls this modulation occurs through alterations of Ca^{2+} flux formed across the sarcoplasmic reticulum (SR) membrane inside cardiomyocytes. Changes in this flux have a profound dependence on the interactions of three proteins: protein kinase A (PKA), sarcoendoplasmic reticulum Ca^{2+}-ATPase (SERCA), and phospholamban (PLN). Phosphorylation of PLN by PKA is associated with an augmented rate of SR Ca^{2+} uptake and relaxation of the myocardium. Mutants of PLN (R9C-PLN and R14Del-PLN) have previously been shown to be linked with forms of the fatal hereditary disease, dilated cardiomyopathy. The molecular basis of disease in this situation could result from irregularities in the association of these PLN mutants with PKA. The work presented here lays the foundation for obtaining the molecular details which govern these interactions to further our understanding of the processes which control Ca^{2+} transport in myocytes and, perhaps, lend insight into the origins of this disease.

Table of Contents

List of Tables

Appendix I

NMR, Mass Spectrometry and Chemical Evidence Reveal a Different Chemical Structure for Methanobactin that Contains Oxazolone Rings

List of Figures

Introduction

Expression and Purification of Isotopically Labeled Peptide Inhibitors and Substrates of cAMP-Dependent Protein Kinase A for NMR Analysis

Carbonyl Carbon Label Selective (CCLS) ^{1}H-^{15}N HSQC Experiment for Improved Detection of Backbone ^{13}C-^{15}N Cross Peaks in Larger Proteins

Simultaneous Detection and Deconvolution of Congested NMR Spectra Containing Three Isotopically Labeled Species

Backbone NMR Resonance Assignment of the Catalytic Subunit of cAMP-Dependent Protein Kinase A in complex with AMP-PNP

Allosteric Cooperativity in Protein Kinase A

Appendix I

NMR, Mass Spectrometry and Chemical Evidence Reveal a Different Chemical Structure for Methanobactin that Contains Oxazolone Rings

Appendix II

Structural and Dynamic Basis of Phospholamban and Sarcolipin Inhibition of Ca^{2+}-ATPase (SERCA)

List of Abbreviations

ATP	Adenosine Triphosphate
ADP	Adenosine Diphosphate
AMP-PNP	aden-5'-yl imidodiphosphate
CCLS	Carbonyl-Carbon Label Selective
DCLS	Dual Carbon Label Selective
CPMG	Carr-Purcell-Meiboom-Gill
DCM	Dilated Cardiomyopathy
DHPR	Dihydropyradine Receptor
DPC	Dodecylphosphocholine
DOPC	1,2-Dioleoyl-*sn*-Glycero-3-Phosphocholine
DOPE	1,2-Dioleoyl-*sn*-Glycero-3-Phosphoethanolamine
ESI-TOF	Electrospray time of flight
HSQC	Heteronuclear Single Quantum Coherence
NMR	Nuclear Magnetic Resonance
NCX	Na^{+}/Ca^{2+} Exchanger
PKA-C	Catalytic Subunit of Protein Kinase A
RyR2	Ryanodine Receptor
SERCA	Sarcoendoplasmic Reticulum Ca-ATPase
ssNMR	Solid State Nuclear Magnetic Resonance
TROSY	Transverse Relaxation Optimized Spectroscopy

Preface

Chapters 2-6 and Appendices I and II are taken directly from manuscripts which are either published, in-press, or submitted.

2. *From the article:* Expression and Purification of Isotopically Labeled Peptide Inhibitors and Substrates of cAMP-Dependent Protein Kinase A for NMR Analysis. Larry R. Masterson, Nadia Bortone, Tao Yu, Kim Ha, Ece C. Gaffarogullari, Oanh Nguyen, and Gianluigi Veglia. (2008*) Protein Expression and Purification*, in-press.

3. *From the article:* Carbonyl Carbon Label Selective (CCLS) ^{1}H-^{15}N HSQC Experiment for Improved Detection of Backbone ^{13}C-^{15}N Cross Peaks in Larger Proteins. Marco Tonelli, Larry R. Masterson, Klaas Hallenga, Gianluigi Veglia, and John L. Markley. (2007) *Journal of Biomolecular NMR* 39, 177-185.

4. *From the article:* Simultaneous Detection and Deconvolution of Congested NMR Spectra Containing Three Isotopically Labeled Species. Larry R. Masterson, Marco Tonelli, John L. Markley, and Gianluigi Veglia. *Journal of the American Chemical Society* (2008) 130, 7818-7819.

5. *From the article:* Backbone NMR Resonance Assignment of the Catalytic Subunit of cAMP-Dependent Protein Kinase A in complex with AMP-PNP. Larry R. Masterson, Lei Shi, Marco Tonelli, Alessandro Mascioni, Michael M. Mueller, and Gianluigi Veglia. Submitted to *Biomolecular NMR Assignments.*

6. *From the article:* Allosteric Cooperativity in Protein Kinase A. Larry R. Masterson, Alessandro Mascioni, Nathaniel J. Traaseth, Susan S. Taylor, and Gianluigi Veglia. *Proceedings of the National Academy of Sciences* (2008) 105, 506-511.

I. *From the article:* NMR, Mass Spectrometry and Chemical Evidence Reveal a Different Chemical Structure for Methanobactin that Contains Oxazolone Rings. Lee A. Behling, Scott C. Hartsel, David E. Lewis, Alan A. DiSpirito, Dong W. Choi, Larry R. Masterson, Gianluigi Veglia, and Warren Gallagher. *Journal of the American Chemical Society* (2008) 130, 12604-12605. Featured in *Chemical and Engineering News*, August 2008.

II. *From the article:* Structural and Dynamic Basis of Phospholamban and Sarcolipin Inhibition of Ca^{2+}-ATPase (SERCA). Nathaniel J. Traaseth, Kim N. Ha, Raffaello Verardi, Lei Shi, Jarrod J. Buffy, Larry R. Masterson, and Gianluigi Veglia. *Biochemistry* (2008) 47, 3-13.

1. Introduction

Physiology and pathophysiology of cardiomyocytes

The reversible phosphorylation of proteins is fundamental to the modulation of myocardial contraction (1). A mechanism which controls this modulation occurs through alterations of Ca^{2+} flux formed across the sarcoplasmic reticulum (SR) membrane inside cardiomyocytes (Figure 1A). Muscle contraction is dependent on the flow of Ca^{2+} into the cytoplasm mediated by the ryanodine receptor (RyR2) from the SR storage, or via the dihydropyradine receptor (DHPR) which shuttles Ca^{2+} at the extracellular membrane. Muscle relaxation has a profound dependence on the interactions of three proteins which remove cytosolic Ca^{2+}: protein kinase A (PKA), sarcoendoplasmic reticulum Ca^{2+} ATPase (SERCA), and phospholamban (PLN). SERCA removes cytosolic Ca^{2+} to initiate relaxation, but the regulatory protein, PLN, decreases SERCA's affinity for Ca^{2+}. When activation of the catalytic-subunit of PKA (PKA-C) occurs during β-adrenergic stimulation, PLN becomes phosphorylated at Ser^{16} by PKA-C (2), which relieves inhibition of SERCA and augments the rate of SR Ca^{2+} uptake. Failure to maintain the balance of this elaborate signaling cascade which controls the translocation of Ca^{2+} results in cardiac disease (3).

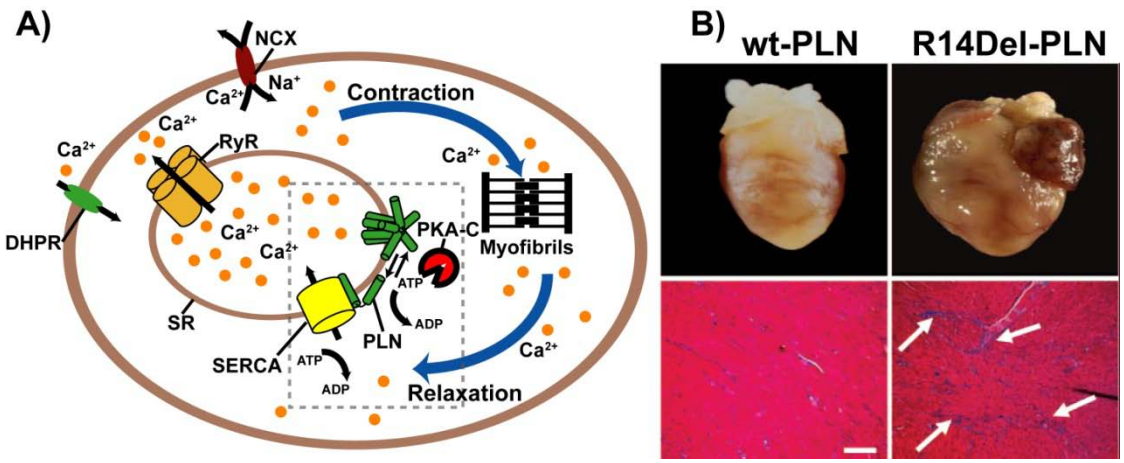


Figure 1: Physiology and pathophysiology of cardiomyocytes. A) Schematic diagram of Ca^{2+} flow within a cardiomyocyte. Influx of extracellular Ca^{2+} from the dihydropyradine receptor (DHPR) and intracellular Ca^{2+} from the sarcoendoplasmic reticulum (SR) into the myocyte cytoplasm induces contraction. Relaxation evolves by the removal of Cytosolic Ca^{2+} by the Na^{+}/Ca^{2+} exchanger (NCX) and, in large part, due to the efflux of Ca^{2+} out of the cell and back into the SR by SERCA which is inhibited by unphosphorylated PLN. Phosphorylation of PLN by PKA-C relieves this inhibition. B) Enlarged heart wall muscle, characteristic of dilated cardiomyopathy, was shown in transgenic mice carrying the R14Del-PLN mutation (upper right panel). Histology of the heart tissue from these mice revealed interstitial fibrosis (lower right panel at arrows). Adopted from reference (4).

Dilated cardiomyopathy (DCM) is an inheritable, fatal disease characterized by an enlargement of the heart, thinning of the heart muscle walls, and weak contractility (Figure 1B). A gain-of-function mutant of PLN, in which the codon for R14 is deleted (R14Del-PLN), has been previously identified in two families with DCM (4). In this study, R14Del-PLN was found to be phosphorylated by PKA-C; however, upon phosphorylation, the inhibition of SERCA was not completely relieved. The hypothesis these authors built based on their experimental data was that the R14Del-PLN mutation caused a shift to more monomeric phospholamban (the inhibitory form) in the pentamer-monomer equilibrium which PLN undergoes (Figure 1A). The repercussion of this genetic mutation is a decrease in the magnitude of relaxation in heart muscle, eventually leading to heart failure and early death in patients which carry it. In one family affected by R14Del-PLN, severe DCM appeared by middle age (4), while in another family, the affected members did not develop DCM until they were in their seventies (5).

A different study (6) established that a missense mutation in a codon for PLN (R→C mutation at residue 9) is also directly correlated to DCM. The data from this study showed that R9C-PLN is unable to properly undergo its regulatory role to inhibit SERCA, and additionally, the R9C-PLN mutant binds irreversibly to PKA-C. Since PKA-C is entrapped, its availability for phosphorylation is decreased. As a consequence, the amount of dephosphorylated wild-type PLN (also expressed in these patients) is increased. Similar to the affects of R14Del-PLN, the repercussion of this genetic mutation was a decrease in the magnitude of relaxation in heart muscle, leading to heart failure and early death in patients carrying this mutation.

An either of the above cases of genetic mutations to PLN, an irregularity in Ca^{2+} handling in cardiomyocytes has been implicated as causation for morbidity. Both mutations lie near the recognition site of PKA-C. Thus, from a molecular perspective, the irregularity in Ca^{2+} handling may be attributed to aberrant protein-protein interactions between the PLN mutants and PKA-C. Understanding the enzymatic interactions of PKA-C with the PLN mutants at an atomic level of detail could provide significant insight into how DCM is manifested in affected patients, and may lead to therapeutic pathways to treat genetically based DCM. Information about the differential interactions of these mutants with PKA-C would need to be unraveled at high-resolution and under conditions which reflect their physiological environment.

PKA: a ubiquitous element of signal transduction

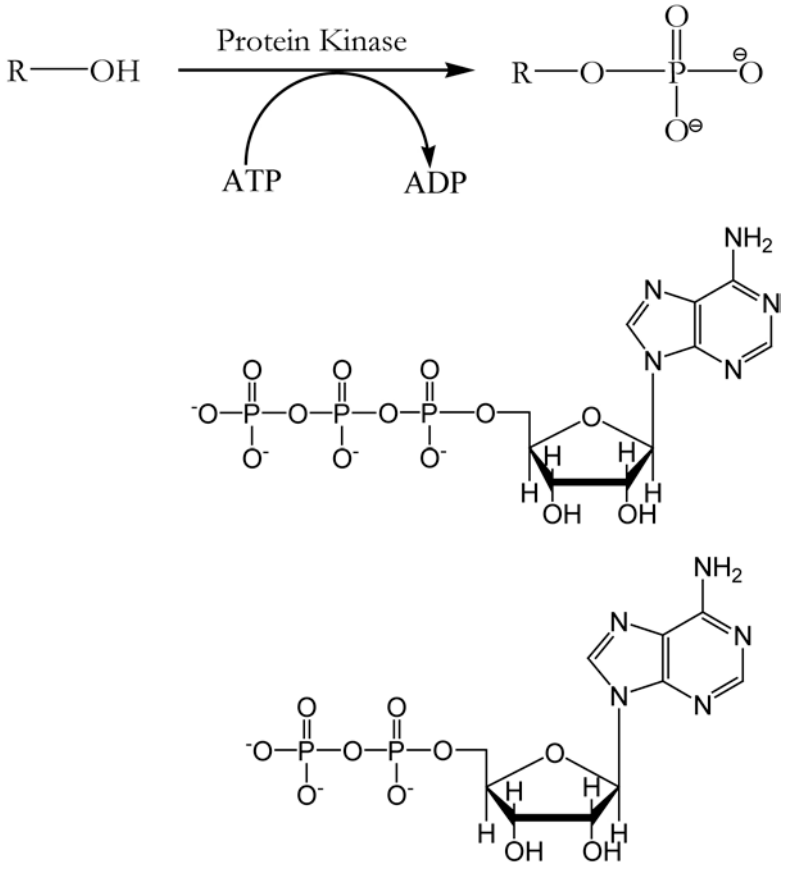
Protein kinases catalyze the transfer of the γ-phosphoryl group of adenosine triphosphate (ATP) to an amino acid hydroxyl group (commonly serine, threonine, or tyrosine) of a substrate (denoted as R in Reaction Scheme 1). The protein kinase family makes up approximately 2000 enzymes in humans, and estimates project that approximately 30% of all proteins encoded by the human genome undergo reversible phosphorylation (7). Abnormal phosphorylation is recognized as a cause or repercussion of numerous diseases such as cancer, diabetes, and rheumatoid arthritis. In fact, the cancer treatment drugs cyclosporine and rapamycin act to inhibit phosphorylation activity (8). Thus, a complete understanding of these reversible phosphorylation events during signal transduction represents perhaps one of the most

important factors in specific understanding of biological function and has vast implications on approaches for the treatment of disease.

Protein kinase A (EC: 2.7.1.37) was the first kinase ever to be studied (9). The catalytic core of this enzyme was found to be the most highly conserved among the family of kinases and has therefore been used as a model enzyme for the other members (10). It exists as a holoenzyme of two regulatory subunits and two catalytic subunits. A conformational change causes the regulatory subunits to release the catalytic subunits upon binding 3', 5'-cyclic adenosine monophosphate (cAMP). PKA-C consists of 350 amino acid residues and can undergo myristoylation to form an amide linked *n*-tetradecanoyl group to its N-terminal glycine via post-translational modification. This modification allows the enzyme to become highly membrane-associated (11, 12) where interactions with membrane proteins can take place.

A wealth of crystallographic data is available on PKA-C, which is a direct result of the ability to express this enzyme as a recombinant protein in *Escherichia coli* (*E. coli*) (13). Well over a dozen high resolution structures have been deposited for the PKA-C in the protein data bank (PDB). These static representations of the enzyme along with fluorescence anisotropy data have revealed that dynamic toggling can occur between three states: an apoenzyme "open state" in which no substrate is bound, transient "intermediate" states, and a "closed" state (14) (Figure 3). Although crystallographic structures are indeed informative, the static models they present are not enough to gain a thorough understanding about the inherently *dynamic nature* of an enzyme such as this. Additionally, interactions with membrane proteins using these methods are not incredibly feasible because the production of high resolution crystals is

Reaction Scheme 1:



difficult in the presence of membranes.

In light of this, high-resolution solution nuclear magnetic resonance (NMR) spectroscopic data of PKA-C has definitive necessity in order that a *dynamic model* can be made of these conformational states in solution. In terms of previous advancements in this regard, we are limited to very little information (15, 16). This limitation is due to the fact that the major step in such analyses requires sequential assignment for the nuclear spins observed in NMR spectra. In fact, one study was only able to achieve a partial backbone assignment (~55%) of the *apo* form of PKA-C (15). The difficulty of obtaining more information for this enzyme are related to its size and presence of conformational dynamics which both tend to broaden NMR signals, making them difficult to detect.

Solution NMR Spectroscopy of Large Biomolecules

Solution NMR spectroscopy has emerged as one of the primary sources of information for high resolution structure and dynamics of biomolecules. The atomic level changes which can be monitored by NMR can only be interpreted once resonance information has been assigned to specific atoms in the protein's primary sequence. For small proteins (<100 residues), the task of assigning the naturally abundant spin ½ ^{1}H nuclei can be achieved using the classical approaches of $^{1}H/^{1}H$ total correlation spectroscopy (TOCSY) in combination with $^{1}H/^{1}H$ nuclear overhauser effect spectroscopy (NOESY) (17). Due to spectral overlap and the difficulty to

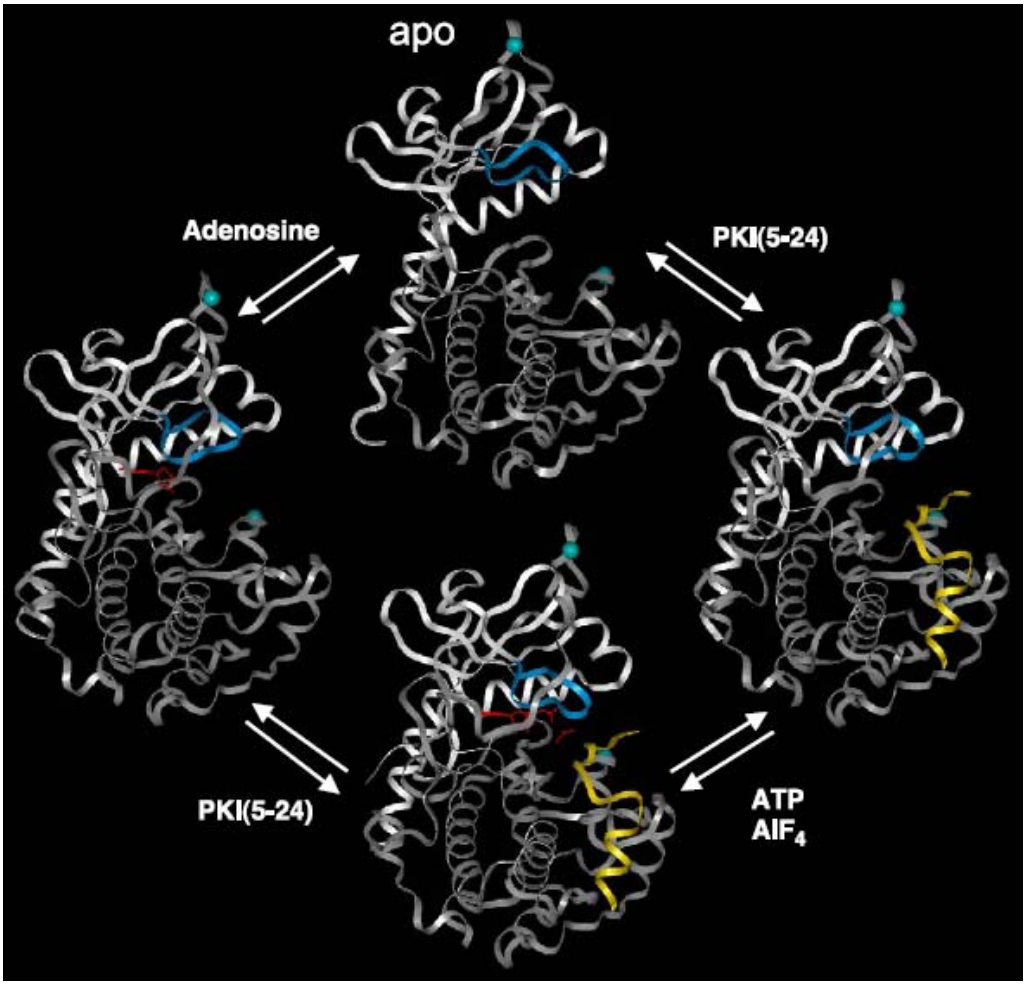


Figure 2: Conformational states of PKA-C. PKA-C has been described to toggle between conformational states, which are selected by ligand binding. Here the crystal structures for the C-subunit enzyme are shown: top is the open apoenzyme containing no substrate or ligand, while the bottom is the closed state. Transitional "intermediate" states have been determined for the enzyme bound to either substrate or ligand. Adopted from (14).

unambiguously assign resonances for medium sized proteins (up to ~25 kDa), isotopic enrichment of other spin ½ nuclei (namely ^{15}N and ^{13}C) must be introduced. This can be achieved in a cost-efficient manner using recombinant proteins expressed in bacterial systems which are grown on glucose and NH_4Cl as sole sources of carbon and nitrogen (18). Higher dimensional (3D and 4D heteronuclear edited) NMR spectroscopy can then be utilized for the purposes of sequential assignment (19). In addition, isotopic enrichment of ^{15}N nuclei allows the acquisition of 2D $^{1}H/^{15}N$ heteronuclear single quantum coherence (HSQC) spectra, which may be thought of as a "fingerprint" correlation spectrum for covalently bonded amide proton and nitrogen nuclei. Since the polypeptide backbone is rich in of these types of correlations, these nuclei are highly sensitive "chemical reporters" for the structural and dynamic changes which can affect the conformation of a protein.

For the cases of large proteins (>25 kDa), inherently short transverse relaxation time values (T_2) of nuclear spins and signal overlap of NMR resonances make detection very difficult (20). This issue becomes worse with increasing molecular weight due to long correlation times, which shorten T_2. However, for large proteins at high magnetic fields (>600 MHz ^{1}H Larmor frequency) two relaxation mechanisms which contribute to T_2 time values of nuclear spins, chemical shift anisotropy and dipole-dipole interaction between amide protons attached to nitrogen nuclei, constructively interfere (21). This narrows the resonance linewidths for one of the four components in an uncoupled $^{1}H/^{15}N$ HSQC spectrum, as shown in Figure 3A. In transverse relaxation optimized spectroscopy (TROSY) (22); this component is exclusively selected for

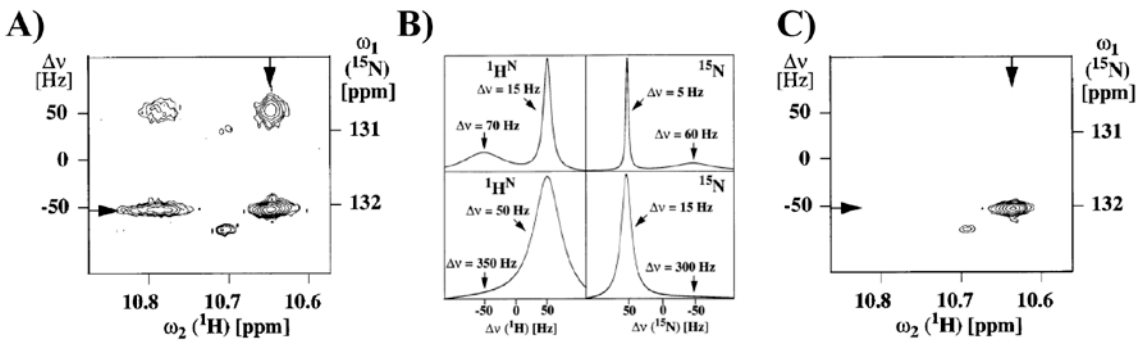


Figure 3: Line broadening encountered in NMR spectroscopy of biomolecules. A) An uncoupled $^{1}H/^{15}N$ HSQC spectrum shows differential linewidths and peak intensities for each of the four components. B) One dimensional traces along either ^{1}H or ^{15}N frequencies for each of these four components reveals a broad and sharp component of each doublet for each dimension. C) TROSY based techniques select for the narrow component of the quartet, containing the most slowly relaxing resonance. This introduces higher sensitivity and improvements of spectral resolution. Adopted from (22).

detection. For large proteins, TROSY affords increased sensitivity and decreased resonance linewidths.

The TROSY effect can be further improved by deuteration of non-exchangeable hydrogen nuclei (20). Deuteration decreases the contribution of dipole-dipole relaxation via surrounding protons (Figure 4). For instance, the major relaxation contribution to carbon magnetization in macromolecules occurs from one-bond dipolar interactions of the directly bonded proton (i.e., ^{1}H-^{13}C) (23). Replacement of ^{1}H with ^{2}H results in the attenuation of this dipolar interaction by a factor of nearly 15. Thus, significant lengthening of carbon relaxation is obtained from deuteration and this allows vast improvements to the sensitivity of experiments which rely on magnetization involving carbon and other spin ½ nuclei (20). In this fashion, TROSY spectroscopy and perdeuteration of large proteins allows access of biomolecules up to 1 MDa (24).

Sequential Assignment of Large Proteins

As previously mentioned, much information can be obtained from the 2D $^{1}H/^{15}N$ HSQC spectrum. The assignment of this spectrum for large proteins, such as PKA-C, is typically achieved via correlations observed from sets of intra- and interresidue chemical shifts of ^{1}H, ^{15}N, and ^{13}C heteronuclei obtained from triple resonance 3D data sets acquired with TROSY detection (Table 1). The chemical shifts of these nuclei are typically sampled in the time domain after transferring magnetization from one nucleus to another through one- or two-bond scalar couplings (Figure 4) and detected on the more sensitive amide proton. Typically, experiments which allow the acquisition of

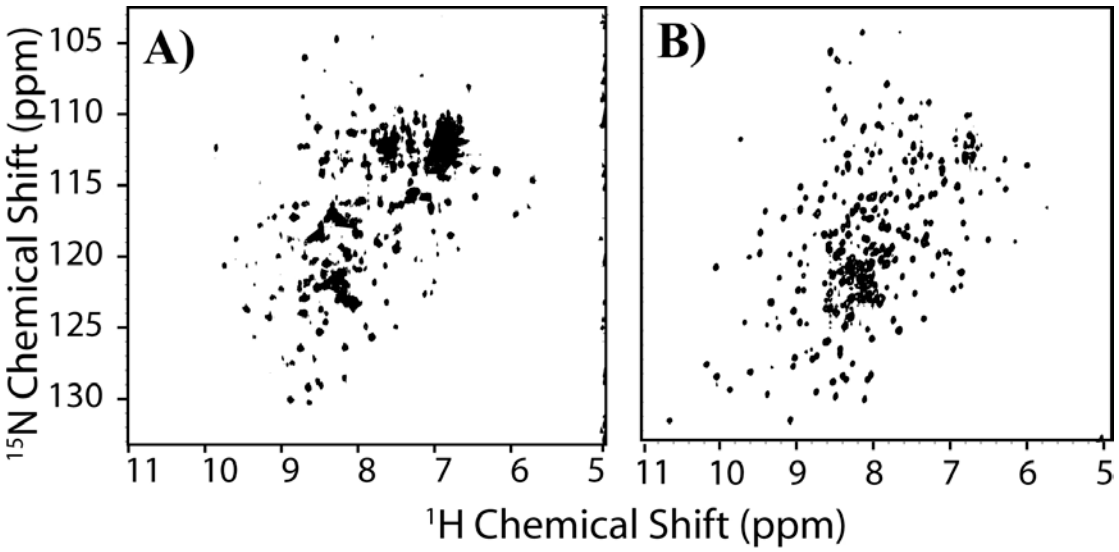
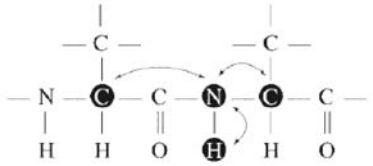
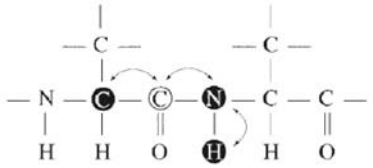
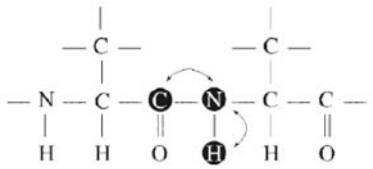
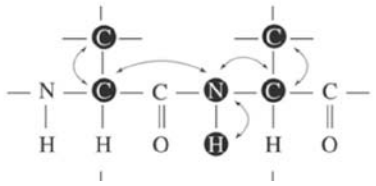


Figure 4: Effects of perdeuteration of large proteins on high resolution NMR spectra. A) PKA-C labeled with ^{15}N and ^{13}C without perdeuteration resulted in the detection of broad amide resonances. The total number of observed peaks in this case was less than 60 percent of the expected resonances for the $^{1}H/^{15}N$-TROSY-HSQC shown here. B) Perdeuterated PKA-C labeled with ^{15}N and ^{13}C showed improved sensitivity and resolution of the NMR resonances in the $^{1}H/^{15}N$-TROSY-HSQC spectrum.

both intra- and interresidue chemical shift correlations in one experiment (i.e., HNCA) will need a complementary experiment in which only interresidue chemical shifts are observed (i.e., HN(CO)CA) in order to make unambiguous connectivities. Given that sufficient sensitivity and resolution are obtained, the information gathered from an experimental data set containing the most sensitive triple resonance experiments for perdeuterated proteins (HNCA, HN(CO)CA, HNCACB, and HNCO) can be sufficient for complete sequential resonance assignment of the protein backbone. However, since residue type information is determined in large part from C^{α} and C^{β} shifts (25), and these frequencies are quite difficult to obtain for large systems, supplementary information is necessary from selective amino acid labeling in which only one residue type is isotopically enriched. This approach works particularly well to acquire 2D $^{1}H/^{15}N$ HSQC spectra of a particular $^{15}N^{\alpha}$ labeled residue type which has a low likelihood biosynthetic exchange in the cellular environment (18). Additional information may also be necessary by using $^{13}C'$ and $^{15}N^{\alpha}$ labeled amino acids to identify the unique dipeptides along the primary sequence of the protein (27) and detection of the unique pairs using a 2D version of the HNCO. Analysis of the data sets from several of these types of triple resonance experiments and selectively labeled samples allows the unambiguous assignment of ^{1}H, ^{15}N, and ^{13}C backbone nuclei, which can therefore be transferred to the assignment of the 2D $^{1}H/^{15}N$ HSQC spectrum.

Table 1: TROSY-based triple resonance experiments used for the backbone assignment of proteins. One-bond couplings are denoted ^{1}J, while two-bond couplings are denoted as ^{2}J. Adopted from (26).

Experiment	Correlation observed	Magnetization transfer	J couplings
3D TROSY-HNCA	$^{1}H_i^{N}$—$^{15}N_i$ —$^{13}C_i^{\alpha}$ $^{1}H_i^{N}$—$^{15}N_i$ —$^{13}C_{i-1}^{\alpha}$		$^{1}J_{NH}$ $^{1}J_{NC^{\alpha}}$ $^{2}J_{NC^{\alpha}}$
3D TROSY-HN(CO)CA	$^{1}H_i^{N}$—$^{15}N_i$ —$^{13}C_{i-1}^{\alpha}$		$^{1}J_{NH}$ $^{1}J_{NCO}$ $^{1}J_{C^{\alpha}CO}$
3D TROSY-HNCO	$^{1}H_i^{N}$—$^{15}N_i$ —$^{13}CO_{i-1}$		$^{1}J_{NH}$ $^{1}J_{NCO}$
3D TROSY-HNCACB	$^{1}H_i^{N}$—$^{15}N_i$ —$^{13}C_i^{\alpha}$/$^{13}C_i^{\beta}$ $^{1}H_i^{N}$—$^{15}N_i$ —$^{13}C_{i-1}^{\alpha}$/$^{13}C_{i-1}^{\beta}$		$^{1}J_{NH}$ $^{1}J_{NC^{\alpha}}$ $^{1}J_{C^{\alpha}C^{\beta}}$ $^{2}J_{NC^{\alpha}}$

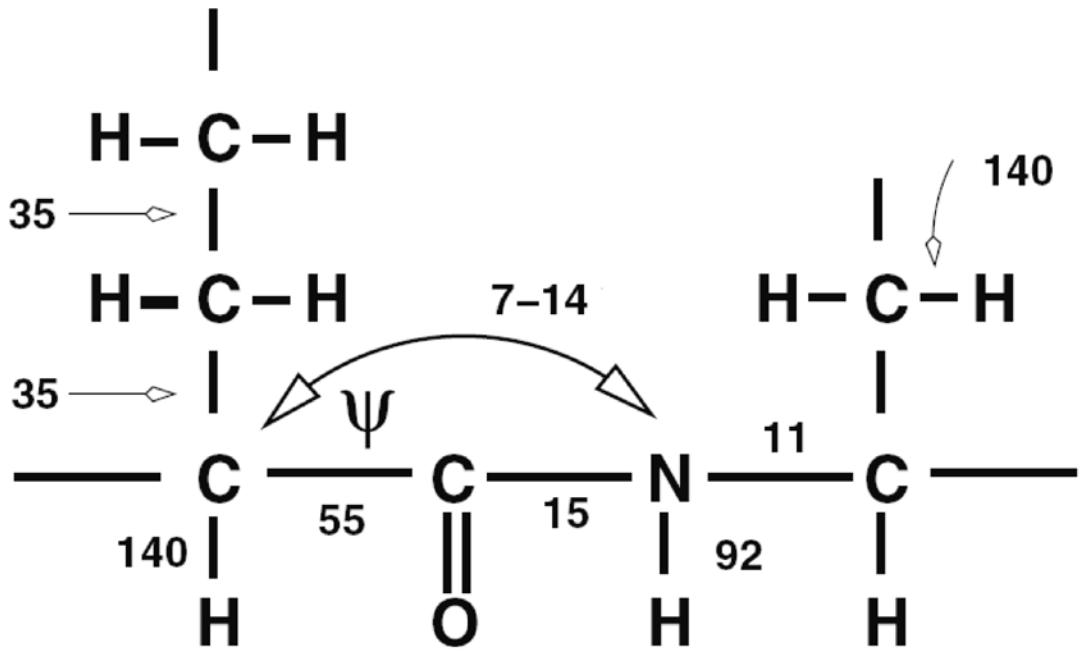


Figure 5: Average scalar coupling constants of ^{1}H, ^{15}N, and ^{13}C nuclei observed along the protein backbone (values are given in Hz). These values will differ depending on the conformation of the backbone and dihedral angles of interacting nuclei. Adopted from (26).

Probing Protein-Protein and Protein-Ligand Interactions

The acquisition of $^{1}H/^{15}N$ HSQC spectra during the titration of an interacting ligand or substrate can be useful for identifying the binding surface of proteins, and in favorable cases, accurate measurements of the thermodynamics and the kinetics which occur during binding can be made (28). The conversion the free protein to a bound state can be viewed simplistically in terms of two-site exchange, going from state A to state B, as in equation 1:

$$A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} B \qquad (1)$$

with forward and reverse rate constant, k_1 and k_{-1}, respectively. A linear combination of these two rates defines the exchange rate, k_{ex}, of this interconversion. Assuming the two states have distinct resonance frequencies (denoted ν_A and ν_B); the type of changes in resonances observed in the spectrum will be determined by the relationship of k_{ex} to the difference of the resonance frequencies. Fast exchange occurs when the difference of frequencies is greater than the exchange rate, intermediate exchange occurs when the difference of frequencies is approximately equal to the exchange rate, and slow exchange occurs when the difference in frequencies is smaller than the exchange rate of interconversion (Figure 5).

Thus, the interconversion or chemical exchange that occurs manifests itself in the observed linewidths detected by NMR spectroscopy. Amide resonances that experience fast exchange phenomena during titrations can be used for quantitative

measurements of the thermodynamic dissociation constant (K_d) for the binding event. The ligand concentration dependent changes in nuclear spin frequency can be used in the following equations using relationships from the Langmuir isotherm (28, 29):

$$\delta_o = \frac{b - \sqrt{b^2 - 4 \times a \times c}}{2a} \quad (1)$$

$$a = \frac{[PKA]_{TOTAL}}{K_d \delta} \quad (2)$$

$$b = 1 + \left(\frac{1}{K_d} \langle [Ligand]_{TOTAL} + [PKA]_{TOTAL} \rangle \right) \quad (3)$$

$$c = \frac{\delta \times [Ligand]_{TOTAL}}{K_d} \quad (4)$$

Where at a given ligand titration point: $[PKA]_{TOTAL}$ is the total concentration of PKA-C present, $[Ligand]_{TOTAL}$ is the total concentration of ligand added, δ is the chemical shift of the nuclei undergoing resonance shift changes, and K_d is the binding constant. Non-linear least squares fitting to obtain the solutions to the quadratic equation 1) can be performed to obtain K_d values for the binding event.

In the limit of slow exchange, the forward and reverse rates of the binding event can be determined from reaction. This can be done using the fact that the following association occurs during binding:

$$A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} B + X \quad (5)$$

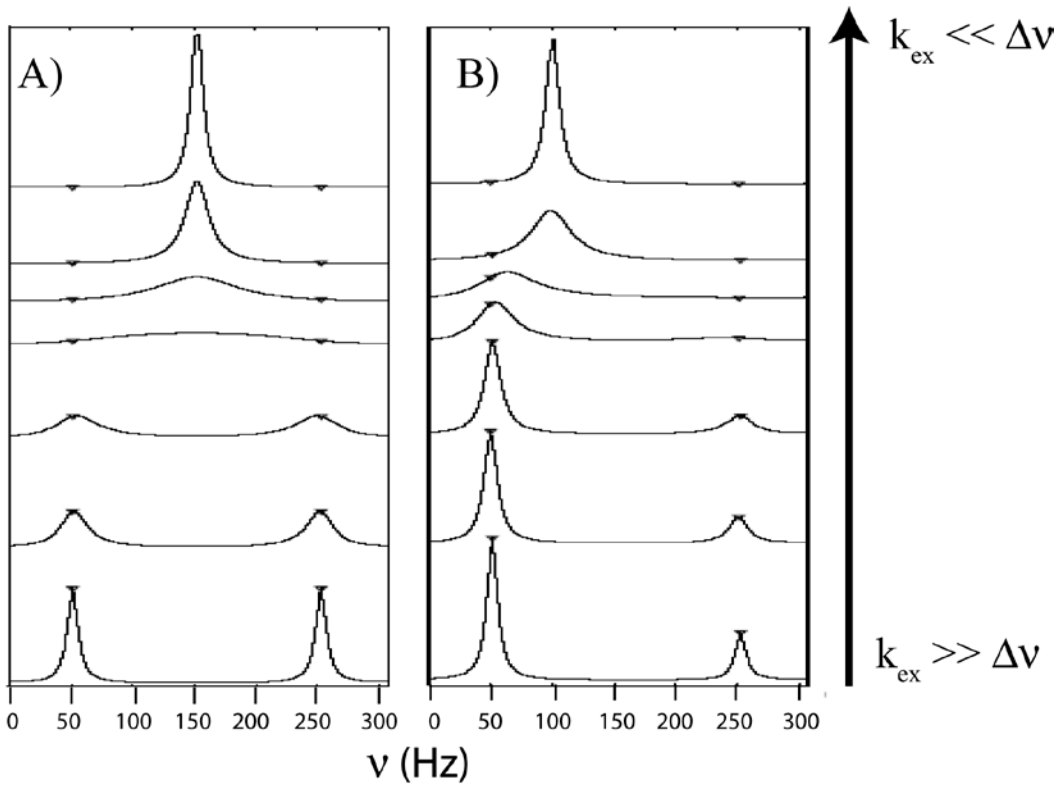


Figure 6: Simulated NMR spectra of a pair of nuclear spins which are related via chemical exchange. Resonances are separated by 200 Hz and the populations of each are either 1:1 (panel A) or 1:0.10 (panel B). When the rate of exchange (k_{ex}) exceeds the difference in Larmor frequencies ($\Delta\nu$), slow exchange occurs. This results in two distinct resonances with intensities proportional to their population. When the rate of exchange is smaller than the $\Delta\nu$, fast exchange occurs. In this case, one resonance is observed at the population weighted average of frequencies. Adopted from (30).

Where k_{-1} is also k_{on} and k_1 is k_{off} for the complex formation. Note here that state B can be thought of as the free protein state. The lifetime of the complex, A, is described by the constant τ_A and the lifetime τ_B is dependent on the concentration of the ligand, X:

$$\frac{1}{\tau_A} = k_{off} \tag{6}$$

$$\frac{1}{\tau_B} = k_{off}[X] \tag{7}$$

Using these constants and relationships to the Bloch-McConnell equations results in a series of the following equations (28, 29):

$$\mathrm{v} = -C_0 \frac{P\left(1 + \tau\left(\frac{\rho_B}{T_{2A}} + \frac{\rho_A}{T_{2B}}\right) + QR\right)}{P^2 + R^2} \tag{8}$$

$$\delta\nu = \nu_A - \nu_B,\ \Delta\nu = \tfrac{1}{2}(\nu_A + \nu_B) - \nu \tag{9}$$

$$P = \tau\left[\frac{1}{T_{2A}T_{2B}} - 4\pi^2\Delta\nu^2 + \pi^2\delta\nu^2\right] + \frac{\rho_A}{T_{2A}} + \frac{\rho_B}{T_{2B}} \tag{10}$$

$$Q = \tau\left[2\pi\Delta\nu - \pi\delta\nu(\rho_A - \rho_B)\right] \tag{11}$$

$$R = 2\pi\Delta\nu\left[1 + \tau\left(\frac{1}{T_{2A}} + \frac{1}{T_{2B}}\right)\right] + \pi\delta\nu\tau\left(\frac{1}{T_{2B}} - \frac{1}{T_{2A}}\right) + \pi\delta\nu(\rho_A - \rho_B) \tag{12}$$

$$\tau = \frac{\rho_A}{k_B} = \frac{\rho_B}{k_A} \tag{13}$$

$$C_0 = \frac{C_A}{\rho_A} = \frac{\gamma_A B_1 M_0}{\rho_A} = \frac{\gamma_A B_1}{\rho_A} \frac{N\mu^2 B_0}{kT} \qquad (14)$$

Where $\Delta\nu$ is the difference in resonance frequency of state A and B, ρ_a and ρ_b are the relative populations of A and B, k is the Boltzmann constant, γ is the gyromagnetic ratio of the nucleus observed, μ is the permeability of free space, T is the temperature, and T_2 is the transverse relaxation time value. A data set containing the changes in frequencies during the titration of free protein to completely bound can be used for lineshape fitting of equation 8 to obtain the on and off-rates.

The changes observed in amide resonances due to an interacting ligand reflect conformational, dynamic, allosteric changes, along with changes in electrostatics, hydrogen bonding, and other factors (33). Residue specific plots of the magnitude chemical shift perturbation therefore reflect any of these changes. Mapping of these perturbations unto the available structural data for PKA-C using combined chemical perturbation at each amide site along the backbone can be done using the following (29, 34):

$$\Delta\delta = \sqrt{[\Delta\delta_H]^2 + [0.154(\Delta\delta_N)]^2} \qquad (15)$$

Where $\Delta\delta$ is the combined chemical shift deviation, and $\Delta\delta_H$ and $\Delta\delta_N$ are the differences in chemical shift between the first and last point of the titration. The weighting factor for the ^{15}N chemical shift is the ratio of the standard deviations of ^1H_N shifts with respect to $^{15}N_H$ shifts averaged over all 20 amino acid types as previously described (35).

References

(1) Mac Lennan DH & Kranias EG (2003) *Nature Reviews* 4, 566-577.

(2) Simmerman HK, Kobayashi YM, Autry JM & Jones LR (1996) *Journal Biological Chemistry* 271, 5941-5946.

(3) Brittsan AG & Kranias EG (2000) *J Mol Cell Cardiol* 32, 2131-9.

(4) Haghighi K, Haghighi K, Kolokathis F, Gramolini AO, Waggoner JR, Pater L, Lynch RA, Fan GC, Tsiapras D, Parekh RR, Dorn GW 2nd, MacLennan DH, Kremastinos DT, Kranias EG.*et al* (2006) *Proc Natl Acad Sci U S A* 103, 1388-1393.

(5) DeWitt MM, MacLeod HM, Soliven B & McNally EM (2006) *J Am Coll Cardiol* 48, 1396-1398.

(6) Schmitt JP, *et al* (2003) *Science* 299, 1410-3.

(7) Cohen P (2002) *Nat Cell Biol* 4, E127-30.

(8) Cohen P (2001) *Eur J Biochem* 268, 5001-5010.

(9) Walsh DA, Perkins JP & Krebs EG (1968) *J Biol Chem* 243, 3763-3765.

(10) Adams JA (2001) *Chem Rev* 101, 2271-2290.

(11) Carr SA, Biemann K, Shoji S, Parmelee DC & Titani K (1982) *Proc Natl Acad Sci U S A* 79, 6128-6131.

(12) Breitenlechner CB, *et al* (2004) *Biochemistry* 43, 7743-7749.

(13) Slice LW & Taylor SS (1989) *J Biol Chem* 264, 20940-20946.

(14) Taylor SS, Yang J, Wu J, Haste NM, Radzio-Andzelm E, Anand G. (2004) PKA: *Biochim Biophys Acta* 1697, 259-269.

(15) Langer T, Vogtherr M, Elshorst B, Betz M, Schieborr U, Saxena K and Schwalbe H (2004) *Chembiochem* 5, 1508-1516.

(16) Langer T, Sreeramulu S, Vogtherr M, Elshorst B, Betz M, Schieborr U, Saxena K, Schwalbe H.*et al* (2005) *FEBS Lett* 579, 4049-4054.

(17) Wuthrich K (1986) *NMR of Proteins and Nucleic Acids.* John Wiley and Sons, New York.

(18) Lian LY & Middleton DA (2001) *Progress in Nuclear Magnetic Resonance Spectroscopy* 39, 171-267.

(19) Bax A & Ikura M (1991) *Journal of Biomolecular NMR* 1, 99-104.

(20) Gardner KH & Kay LE (1998) *Annu Rev Biophys Biomol Struct* 27, 357-406.

(21) Dalvit C (1992) *Journal of Magnetic Resonance (1969)* 97, 645-650.

(22) Pervushin K, Riek R, Wider G & Wuthrich K (1997) *Proc Natl Acad Sci U S A* 94, 12366-71.

(23) Browne DT, Kenyon GL, Packer EL, Sternlicht H & Wilson DM (1973) *J Am Chem Soc* 95, 1316-1323.

(24) Fiaux J, Bertelsen EB, Horwich AL & Wuthrich K (2002) *Nature* 418, 207-211.

(25) Wang Y & Jardetzky O (2002) *Protein Sci* 11, 852-861.

(26) Downing AK (2004) *Protein NMR techniques*. Human Press, New Jersey.

(27) Kainosho M & Tsuji T (1982) *Biochemistry* 21, 6273-6279.

(28) Johnson PE, Tomme P, Joshi MD & McIntosh LP (1996) *Biochemistry* 35, 13895-13906.

(29) Tugarinov V & Kay LE (2003) *J Mol Biol* 327, 1121-33.

(30) Abergel D & Palmer AG (2004) *Chemphyschem* 5, 787-793.

(31) Hansen DF & Led JJ (2003) *J Magn Reson* 163, 215-227.

(32) Tugarinov V, Hwang PM & Kay LE (2004) *Annu Rev Biochem* 73, 107-46.

(33) Sitkoff D & Case DA (1997) *J Am Chem Soc* 119, 12262-12273.

(34) Revington M, Holder TM & Zuiderweg ER (2004) *J Biol Chem* 279, 33958-33967.

(35) Evenas J,Evenäs J, Tugarinov V, Skrynnikov NR, . Goto NK, Muhandiram R, Kay LE (2001) *J Mol Biol* 309, 961-974.

2. Expression and Purification of Isotopically Labeled Peptide Inhibitors and Substrates of cAMP-Dependent Protein Kinase A for NMR Analysis

Larry R. Masterson, Nadia Bortone, Tao Yu, Kim Ha, Ece C. Gaffarogullari, Oanh Nguyen, and Gianluigi Veglia

Reprinted with permission from Protein Expression and Purification (2008)

Extensive x-ray crystallographic studies carried out on the catalytic-subunit of protein kinase A (PKAc) enabled the atomic characterization of inhibitor and/or substrate peptide analogs trapped at its active site. Yet, the structural and dynamic transitions of these peptides from the free to the bound state are missing. These conformational transitions are central to understanding molecular recognition and enzymatic cycle. NMR allows one to study these phenomena under functionally relevant conditions. However, the amounts of isotopically labeled peptides required for this technique present prohibitive costs for peptide synthesis. To enable NMR studies, we have optimized both expression and purification of isotopically enriched substrate/inhibitor peptides using a recombinant fusion protein system. Three of these peptides corresponded to the cytoplasmic regions of the wild-type and lethal mutants of the membrane protein phospholamban, while the fourth peptide corresponded to the binding epitope of the heat-stable protein kinase inhibitor (PKI_{5-24}). The target peptides were fused to the maltose binding protein (MBP), which is further purified using a His_6 tag approach. This convenient protocol allows for the purification of milligram amounts of peptides necessary for NMR analysis.

Introduction

Phosphorylation of protein substrates containing a serine or threonine residue by the catalytic subunit of cAMP-dependent protein kinase A (PKAc, EC 2.7.11.11) is a ubiquitous step in the regulation of many metabolic pathways (1). Substrates for this enzyme contain the consensus sequence Arg-Arg-X-Ser/Thr-Y, where X can be any residue and Y is typically a large hydrophobic residue (2). Much of our understanding for the atomic details of PKAc-substrate interactions has been guided by crystallographic data of inhibitor or pseudo-substrate bound complexes (3). These data are focused on analogs of the high affinity binding region (residues 5 to 24) of the heat-stable protein kinase inhibitor (PKI). Although these crystallographic snapshots have provided a wealth of information, the biological relevance obtained is somewhat limited and dynamic information is absent. Obtaining both high resolution structure and dynamic information for this system is important since structure and dynamics are hypothesized to play fundamental roles in enzymatic function (3), but they are difficult to study under functional conditions. Acquiring this type of data is even more difficult for the case of membrane bound or membrane associated protein substrates.

Nuclear magnetic resonance (NMR) provides an avenue to acquire structural and dynamic information of biologically relevant protein substrates or inhibitors of PKAc, although such studies would require milligram quantities of isotopically labeled proteins. In this respect, NMR analysis has been limited to a structure determination performed on full-length PKI and two shorter PKI constructs corresponding to the high affinity PKAc binding region (PKI_{1-24}) and the nuclear export factor (PKI_{26-75}) (4). This

study provided the solution structures of these proteins, but did not provide insight about the changes in dynamics in the absence or presence of PKAc. Moreover, the analysis of PKI_{1-25} at natural abundance restricted the spectral resolution needed to provide the atomic detail for dynamic changes, echoing the need for a method to isotopically enrich NMR active nuclei for this short peptide.

The natural PKAc protein substrate, phospholamban (PLN), could provide more biologically relevant information about how structural dynamics of substrates change upon PKAc interaction. PLN is an integral membrane protein in the sarcoplasmic reticulum of cardiomyocytes where it inhibits the Ca^{2+}ATPase, SERCA (5). NMR analyses of PLN by our laboratory have shown that the structure of this protein contains both a transmembrane and a cytoplasmic helix linked by a short, flexible loop (6, 7). The cytoplasmic region of PLN contains a consensus sequence for PKAc-catalyzed phosphorylation which, in turn, relieves the inhibition of SERCA (8). The cytoplasmic region of PLN was also found to contain two types of mutations which are correlated with dilated cardiomyopathy: an arginine to cysteine mutation at position 9 (PLN-R9C) (9) and a deletion of arginine 14 (PLN-R14Del) (10). Since both mutations are in the vicinity of the PKAc binding region, they could result in a change in interaction with PKAc.

NMR will be the method of choice to understand how structural dynamics are affected by these mutants during substrate recognition and phosphorylation by PKAc at the atomic level. A major obstacle for such studies is that a suitable membrane mimicking environment must be compatible for both proteins and finding the correct conditions can be quite challenging (11). However, model peptides corresponding to

the cytoplasmic residues 1-20 of PLN, PLN-R9C, and PLN-R14Del could be used instead to provide such insight. Since much more is known about PKI_{5-24} from crystallography, comparisons could also be made with this peptide based upon further high resolution NMR analyses of isotopically enriched peptides.

Any of these peptides could be obtained by solid-phase peptide synthesis (SSPS), although it does not permit a cost-efficient approach for peptides isotopically enriched with ^{15}N, ^{13}C, or ^{2}H. Our aim here is to devise a recombinant system to express four different peptide constructs: the cytoplasmic region of PLN (PLN_{1-20}) and its two mutants, R9C (R9C-PLN_{1-20}) and R14Del (R14del-PLN_{1-20}), as well as the high affinity region of the PKAc inhibitor, PKI_{5-24}. We used *Escherichia coli* (*E. coli*) BL21(DE3) cells to overexpress fusion proteins composed of maltose-binding protein (MBP) and a target peptide, separated by a TEV cleavable linker region. As noted for other difficult recombinant proteins (12, 13), this fusion system combined increased resistance to proteolytic degradation of the peptide in the host cell with the ability to obtain milligram quantities of isotopically labeled peptide per liter of media. The introduction of a His_6 tag and a TEV cleavage site to the C-terminus of MBP also allowed a convenient approach to obtain these peptides at purities which exceed 95%.

Experimental Procedures

Materials

All oligonucleotide synthesis and DNA sequencing were performed at the BioMedical Genomics Center of the University of Minnesota. All fusion constructs were cloned using the expression plasmid pMal-c2e (New England Biolabs). The

plasmid encoding the tobacco etch virus protease containing a His_6 tag (TEV) used for enzymatic cleavage of fusion protein was provided as a kind gift from Dr. Robert Gorelick at the National Institute of Health. Expression and purification of TEV was performed as previously described (14). The *E. coli* strain XL1-Blue (Stratagene) was used for plasmid cloning, while strain BL21(DE3) (Novagen) was used for protein expression. Ni^{2+}-NTA resin (HIS-Select^{TM} Nickel Affinity Gel) was obtained from Sigma.

Cloning of peptide inhibitors and substrates

The gene for PLN_{1-20} was designed using the wild type PLN parental template which contained codons optimized for usage in *E. coli* (12). An *Eco*RI site was encoded at the 5'-end, followed by a His_6-tag and a TEV protease cleavage site and the sequence for the peptide of interest. A *Hind*III site was encoded at the 3'-end. The His_6-TEV-PLN_{1-20} gene was amplified using the forward primer 5'-CCG GAA TTC CAT CAT CAT CAT CAT CAT GAA AAC CTG TAT TTT CAG GGC ATG GAA AAA GTG - 3'. The gene for PKI_{5-24} was also designed from oligonucleotides optimized for usage in *E. coli* (4). Similar upstream elements were introduced to this gene as the PLN peptide constructs. Modifications of the PLN_{1-20} sequence to produce the mutant analogue peptides were performed using the Stratagene Quickchange Kit.

Polymerase chain reaction (PCR) was performed using 1X buffer, forward and reverse primers (0.5 μM), dNTP mix (200 μM), Pfu Turbo polymerase (Stratagene; 5 U), and ddH_2O to final volume of 100 μl. The double-stranded DNA was amplified with a first cycle of melting at 94 °C for 2.5 min, elongation at 55 °C for one min, and

annealing at 72 °C for one min; and 29 more cycles with melting at 94 °C for one min, elongation at 60 °C for one min, and annealing at 72 °C for one min. The double stranded DNA containing a target peptide construct (His_6-TEV with either PLN_{1-20}, R9C-PLN_{1-20}, R14del-PLN_{1-20}, or PKI_{5-24}) and pMal-c2e plasmid were digested with *Eco*RI and *Hind*III restriction enzymes, and purified by agarose gel. The digested products were ligated back into the pMal-c2e vector using T4 DNA ligase (Invitrogen) and transformed into XL1-Blue *E. coli* competent cloning cells. DNA purification was performed with the Quick-Spin Miniprep kit (Qiagen) and quantitated by measuring UV absorption at 260 nm. Correct PCR products were confirmed with DNA sequencing. Plasmids encoding for MBP-His_6-TEV-peptide were then transformed into *E. coli* strain BL21(DE3) competent cells.

Expression of fusion proteins

A single colony containing the plasmid for MBP-His_6-TEV-peptide was inoculated into 1 L of sterile Luria-Bertani (LB) medium with 1 mM ampicillin and incubated with shaking at 30 °C. After reaching an OD_{600} of 1.2 (~12 hours), the culture was centrifuged at 3,000*g* for 10 min at room temperature. The pellet was resuspended in 2.5 L of M9 minimal media containing 1 mM ampicillin and incubated with shaking at 37 °C. Once the cells reached an OD_{600} of ~0.9, they were induced with 1 mM IPTG. The cells were allowed to express for 5 hours and were then harvested by centrifugation at 6,000*g* for 20 min at 4 °C. Approximately 10-14 grams of cells (weight mass) were typically obtained from 2.5 L of media. The cell pellet was collected, flash frozen in liquid nitrogen, and stored at -20 ^{o}C.

Cell lysis

Frozen cell pellets were resuspended in 100 ml lysis buffer (0.1 M sodium phosphate, pH 8.0, 6 M guanidinium hydrochloride, 10 mM imidazole), and homogenized in a blender for 10 minutes on ice. The lysis mixture was sonicated on ice with a probe sonicator (Branson Sonifier 450) at 40% duty cycle and output control of 4 for 10 min. Cell debris was cleared by centrifugation at 45,000*g* for 20 min at 4 °C. The supernatant containing fusion protein was collected for purification.

Protein purification

The supernatant was bound batch-wise to a slurry containing 50 ml of Ni^{2+}-NTA resin at room temperature for 20 min with stirring. The protein/resin mixture was loaded onto a column and the flow through was collected. The resin was subsequently washed twice with 50 ml of lysis buffer. The fusion protein was then isolated by washing the resin with lysis buffer containing 50 mM imidazole until UV absorption at 280 nm was less than 0.1 (~150 ml). Fractions containing isolated fusion protein were confirmed via 16 % SDS-PAGE and coomassie staining. The fractions were collected and dialyzed using a 10 kDa molecular weight cut-off membrane at 4 ºC for 2 hours against 3 L of Tris buffer (100 mM Tris, pH 8, 1 M urea, 1 % glycerol), followed by an additional 3 hours at 4 ºC in the same buffer without urea.

Following dialysis, the fusion protein was pooled and enzymatic cleavage was initiated by addition of 1 mg TEV protease/40 mg fusion and stirred by shaking at 30° C for 6 hours. The reaction was stopped by fast unfolding with the addition of

guanidinium hydrochloride to a final concentration of 3.0 M and the solution was bound batch-wise to a slurry containing 50 ml of Ni^{2+} resin at room temperature for 20 min with stirring. The protein/resin mixture was loaded onto a column and the flow through was collected along with a single 25 ml wash. The solution was filtered through a 0.2 μm membrane and loaded onto a semi-preparative reverse phase C-18 Vydac HPLC cartridge (2.5 x 10 cm, 15 μm, 300 Å). Loading was ~30 mg total protein per injection based on $\varepsilon_{MBP} = 1.6\ mg^{-1}*cm^{-1}$ at 280 nm. Peptides were eluted using a linear gradient from Buffer A (0.1% TFA in H_2O) to 40% Buffer B (0.1% TFA in CH_3CN) and detected at 220 nm. Purities of the peptides exceeded 95% as determined by analytical HPLC using a Waters C18 column (0.46 x 25 cm) of the collected peptide peak. Pooled fractions containing peptide were lyophilized immediately after purification.

Quantification and protein identification

Protein concentrations prior to TEV cleavage was determined by absorbance at 280 nm or via gel densitometry using a BioRad Molecular Imager FX using BioRad Quantity One software. Final concentrations of pure peptides were determined using analytical HPLC and standard addition of synthetic peptides which were quantified by amino acid analysis. Pure recombinant peptides were confirmed by ESI-TOF: PKI_{5-24} calculated 2279.4 *m/z*, found 2278.6 *m/z* ; PLN_{1-20} calculated 2440.8 *m/z*, found 2440.1 *m/z*; R9C-PLN_{1-20} calculated 2387.9 *m/z*, found 2387.1 *m/z*; R14del-PLN_{1-20} calculated 2284.4 *m/z*, found 2283.7 *m/z*.

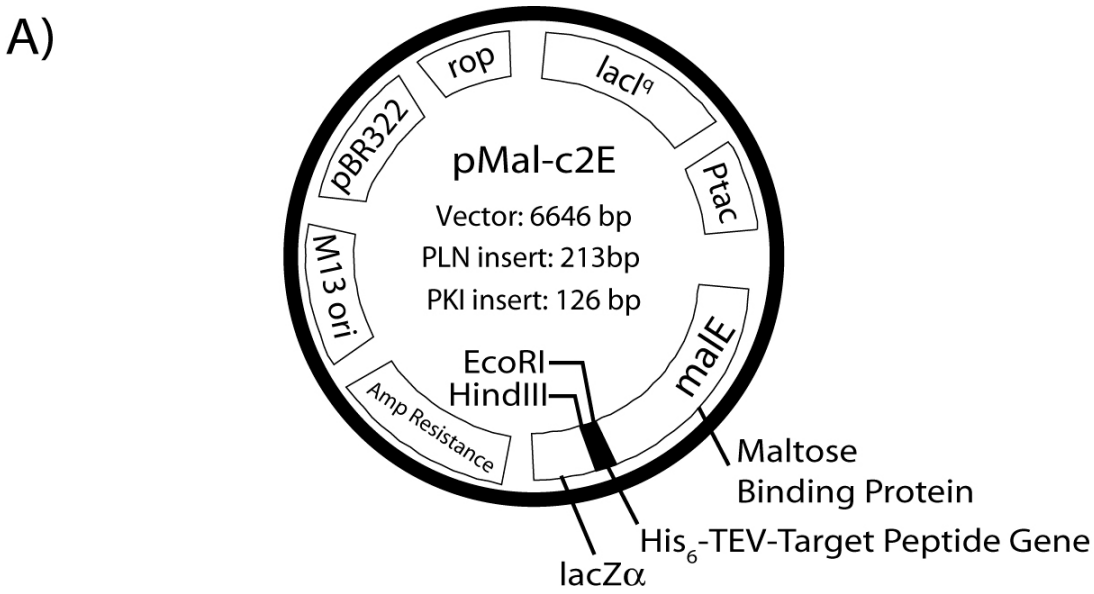
NMR spectroscopy

All peptides were dissolved to a concentration of 1 mM in 20 mM phosphate buffer (pH 6.5), 150 mM KCl, 1 mM NaN_3, 5 mM DTT, and 5% D_2O. $^{1}H/^{15}N$ heteronuclear single quantum coherence (HSQC) spectra were acquired at 300 K on a Varian Inova 600 MHz spectrometer equipped with a triple resonance probe with z-axis gradients. Spectra were processed using NMRPipe (15) and visualized with the software SPARKY (16).

Results and Discussion

Plasmid construct and protein expression

The general plasmid construct and primary sequences of the target peptides studied here are shown in Figure 1. Previous results published by our laboratory for the production of PLN and sarcolipin showed that the selection of MBP as a soluble fusion partner results in milligram quantities of these proteins (12). We extended that approach to obtain these short recombinant peptide analogues of the PLN cytoplasmic region, and the high affinity binding epitope of PKI. One difference of the construct used here is that a modification of the MBP primary sequence was done by introducing a C-terminal His_6 tag (Figure 1) before the TEV cleavage site. This provided a robust approach to isolate the fusion protein from the cell lysate via Ni^{2+}-NTA chromatography. The introduction of this tag was also useful in latter purification steps to allow removal of MBP and TEV protease (which also contained a His_6 tag) after fusion protein cleavage. Yields for MBP fusion constructs were typically 80-125 mg/L



B)

	His$_6$	TEV	Target Protein
PKI$_{5-24}$	HHHHHH	ENLYFQ	GTTYADFIASGRTGRRNAIHD
PLN$_{1-20}$	HHHHHH	ENLYFQ	GMEKVQYLTRSAIRRASTIEM
R9C-PLN$_{1-20}$	HHHHHH	ENLYFQ	GMEKVQYLTCSAIRRASTIEM
R14del-PLN$_{1-20}$	HHHHHH	ENLYFQ	GMEKVQYLTRSAIRASTIEM

Figure 1: Architecture of the plasmid constructs used for the expression of recombinant peptides. A) Plasmid constructs used for the production of recombinant peptides expressed as a fusion protein with MBP. The *Eco*RI and *Hind*III endonuclease sites were used to introduce a DNA insert containing a His_6-(TEV cleavage site)-peptide codon sequence which was optimized for expression in *E. coli*. B) The primary amino acid sequences which were fused to the C-terminus of MBP. Due to the introduction of a TEV protease site (shown simply as TEV above the six residue sequence), each construct is expressed with an additional N-terminal glycine.

Table 1

Protein yields at various purification steps for PKI5-24, PLN1-20, R14del-PLN1-20, and R9C-PLN1-20 recombinant peptides.

Protein	**Amount (mg/L)**	**Purity (%)**	**Fold enrichment**[a]
E. coli soluble fraction (MBP/PKI_{5-24} fusion)[b]	90	6	-
Ni^{2+}-NTA chromatography (MBP/PKI_{5-24} fusion)[b]	90	>90	15
HPLC chromatography (isolated PKI_{5-24})[c]	3	>95	1.1
E. coli soluble fraction (MBP/PLN_{1-20} or R14del-PLN_{1-20}fusion)[b]	75	4	-
Ni^{2+}-NTA chromatography (MBP/PLN_{1-20} or R14del-PLN_{1-20}fusion)[b]	75	>90	23
HPLC chromatography (isolated PLN_{1-20} or R14del-PLN_{1-20})[c]	2	>95	1.1
E. coli soluble fraction (MBP/PLN-$R9C_{1-20}$ fusion)[b]	60	2	-
Ni^{2+}-NTA chromatography (MBP/R9C-PLN_{1-20} fusion)[b]	60	>90	45
HPLC chromatography (isolated R9C-PLN_{1-20})[c]	1	>95	1.1

[a] Fold enrichment is calculated as the percent purity of purification step divided by percent purity of preceding step.

[b] Gel densitometric measurements of 14% SDS-PAGE were used to estimate protein quantities and purity.

[c] Analytical HPLC of the isolated peptide was used to calculate final purity.

in LB or 60-90 mg/L in M9 minimal media with >90% purity by SDS-PAGE (Table 1). A similar approach has been successfully applied to the expression and purification of isotopically labeled transmembrane domains by Cross and co-workers (13).

Protein purification

The scheme used to obtain high purity recombinant peptides in is shown in Figure 2. One key procedure in this scheme is the unfolding step to fully expose the C-terminal His_6-tag between MBP and the peptide to bind Ni^{2+}-NTA resin, followed by a refolding step to allow TEV cleavage to occur efficiently. Isolation of the fusion protein from the cell lysate was done in a batch-wise manner under denaturing conditions. Initial attempts to unfold MBP in the presence of 8 M urea were successful, but a small amount of peptide contaminant 44 amu larger as determined by ESI-TOF mass spectroscopy was consistently observed, regardless of the construct purified. This increased mass corresponds to a carbamylated amine group (either at the N-terminus or at an amine bearing sidechain, such as lysine) via isocyanate (17). Urea is in equilibrium with ammonium and isocyanate and was the likely source for this amine modification. We decided to use 6 M guanidine-hydrochloride to unfold the fusion protein. After isolation of the fusion protein, protein mis-folding and precipitation tended to occur when guanidine-HCl was dialyzed from the solution. To prevent this from occurring, a short exposure to 1 M urea (two hours) in the presence of one percent glycerol at 4 ºC was performed during a first refolding step, followed by a second exposure to one percent glycerol in Tris buffer. The short exposure to urea did not result in any detectable levels of carbamylated product after HPLC purification.

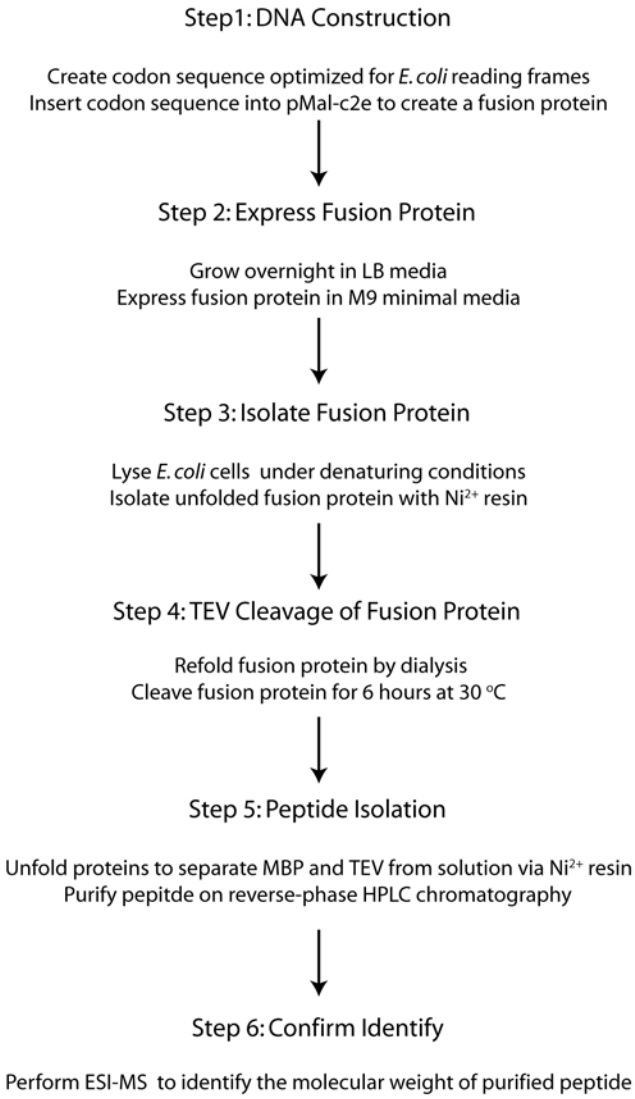


Figure 2: Flowchart for the expression and purification of the recombinant peptides studied here.

TEV cleavage of these fusion proteins was optimal with ~1 mg of TEV protease per 40 mg of fusion protein at 30 ºC. In contrast to previously published conditions for this step, use of Triton X-100 to increase protein solubility during cleavage was avoided. Triton X-100 contains trace amounts of ethylene oxide (18) which can form polyethylene glycol (PEG) in water. Trace amounts of PEG co-eluted with these peptides during HPLC purification and obscured proper identification of the target peptides by mass spectrometry analysis, showing a distinct polymeric pattern separated by 44 amu. Therefore, the concentration of the fusion protein was kept below 1 mg/ml with one percent glycerol to increase solubility of all cleavage products which resulted during reactions. Fusion protein cleavage was >90% complete in 6 hours for all constructs as determined by the differential migration of fusion protein and MBP-His_6-TEV cleavage product (~2.5 kDa smaller) using a 16% SDS-PAGE gel (Figure 3, indicated by at the arrows). Under the dilute conditions used at this step, the cleaved peptides could not be visualized on SDS-PAGE gels using either Coomassie blue dye or silver stain. Cleavage reactions were stopped by rapid unfolding of the solution using guanidine-HCl to a final concentration of 3 M. This allowed efficient extraction of the bulk of MBP and TEV from the reaction solution via Ni^{2+}-NTA chromatography (Figure 3) and well resolved peptide elution during reverse-phase HPLC (Figure 4).

Typically after the cleavage reactions, the cleaved fusion protein did not bind completely to Ni^{2+}-NTA resin, which resulted in small amounts carried to HPLC purification. However, as shown in the representative chromatogram of R14del-PLN_{1-20} in Figure 4, all recombinant peptides were well resolved from any other contaminant from the fusion cleavage step reverse-phase chromatography. A representative ESI-TOF

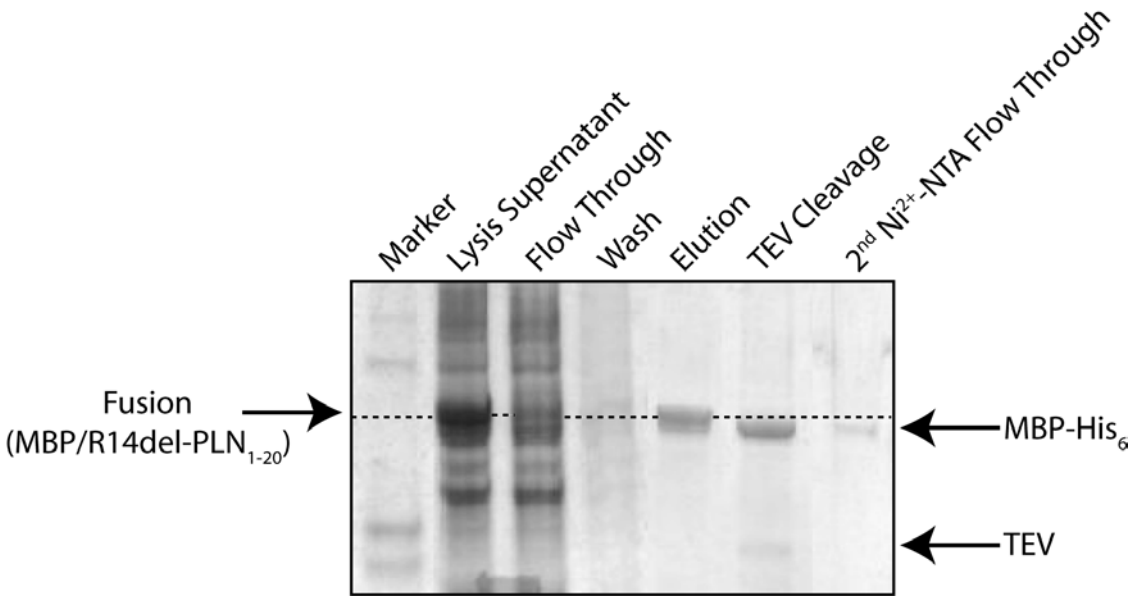


Figure 3: Sixteen percent SDS-PAGE gel depicting the purification of R14del-PLN1-20 peptide starting from selection via Ni^{2+}-NTA chromatography and ending with TEV protease cleavage of the fusion protein. Due to the lack of ability to bind coomassie blue dye or silver stain, the peptide was not detected. The arrow on the left corresponds to the full fusion protein, MBP-His_6-TEV-R14del-PLN_{1-20}, while the arrows at the right correspond to cleaved fusion (MBP-His_6-TEV) and TEV protease. The small difference in migration (indicated by the dashed line) before and after cleavage reactions was used to monitor cleavage reactions. After the cleavage reaction, residual cleaved fusion protein which did not bind efficiently to the Ni^{2+}-NTA resin was typically found in the 2^{nd} flow. These results also reflect those obtained for the three other recombinant peptides.

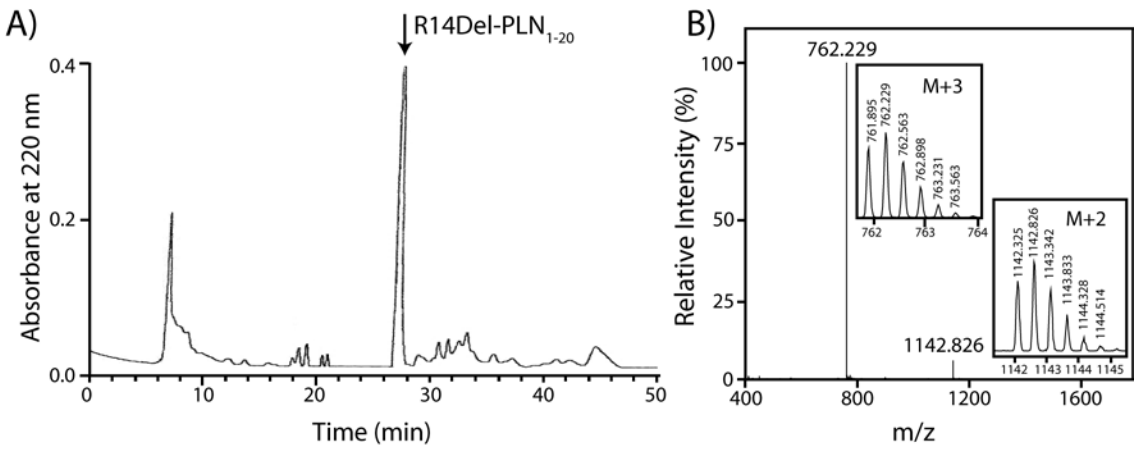


Figure 4: Isolation and identification of cleaved peptides. A) Chromatogram of the semi-preparative reverse-phase HPLC of R14del-PLN_{1-20} peptide after TEV cleavage from the MBP fusion partner. The peak shown at the arrow corresponds to pure R14del-PLN_{1-20}. B) ESI-TOF mass spectroscopy of HPLC purified R14del-PLN_{1-20}. The two major peaks which were detected correspond to M+3 and M+2 positive ionization states of the peptide and respective patterns are shown in the insets.

spectrum of the pooled fractions containing R14del-PLN_{1-20} obtained from HPLC purification is shown in Figure 4.

NMR Spectroscopy

The amide fingerprint of ^{15}N-labeled recombinant peptides acquired using $^{1}H/^{15}N$ HSQC NMR spectroscopy under similar conditions to those used to study PKAc by our laboratory (19) is shown in Figure 5. The amide resonance linewidths displayed by all of these spectra were indicative of peptides which are not aggregated, while the lack of dispersion in the ^{15}N and ^{1}H dimensions for the PLN peptide analogues suggested that these peptides adopt predominant random coil conformations in aqueous buffer. On the other hand, the more dispersed $^{1}H/^{15}N$ HSQC spectrum for PKI_{5-25} may be indicative that this peptide assumes a more defined secondary structure. The latter agrees well with homonuclear solution NMR data previously published (4). From these HSQC spectra, PKI_{5-24} displayed 18 of 20 expected peaks (two peaks contained overlapping resonances, Figure 5A), while PLN_{1-20} and R9C-PLN_{1-20} showed 17 of 20 expected peaks (each had three peaks containing overlapping resonances, Figure 5B and 5C), and R14del-PLN_{1-20} displayed 18 of 19 expected peaks (one peak contained overlapping resonances, Figure 5D).

Conclusions

We report a robust method to express and isolate high purity recombinant peptides (~20 amino acids), corresponding to the PKAc binding regions of PKI, PLN

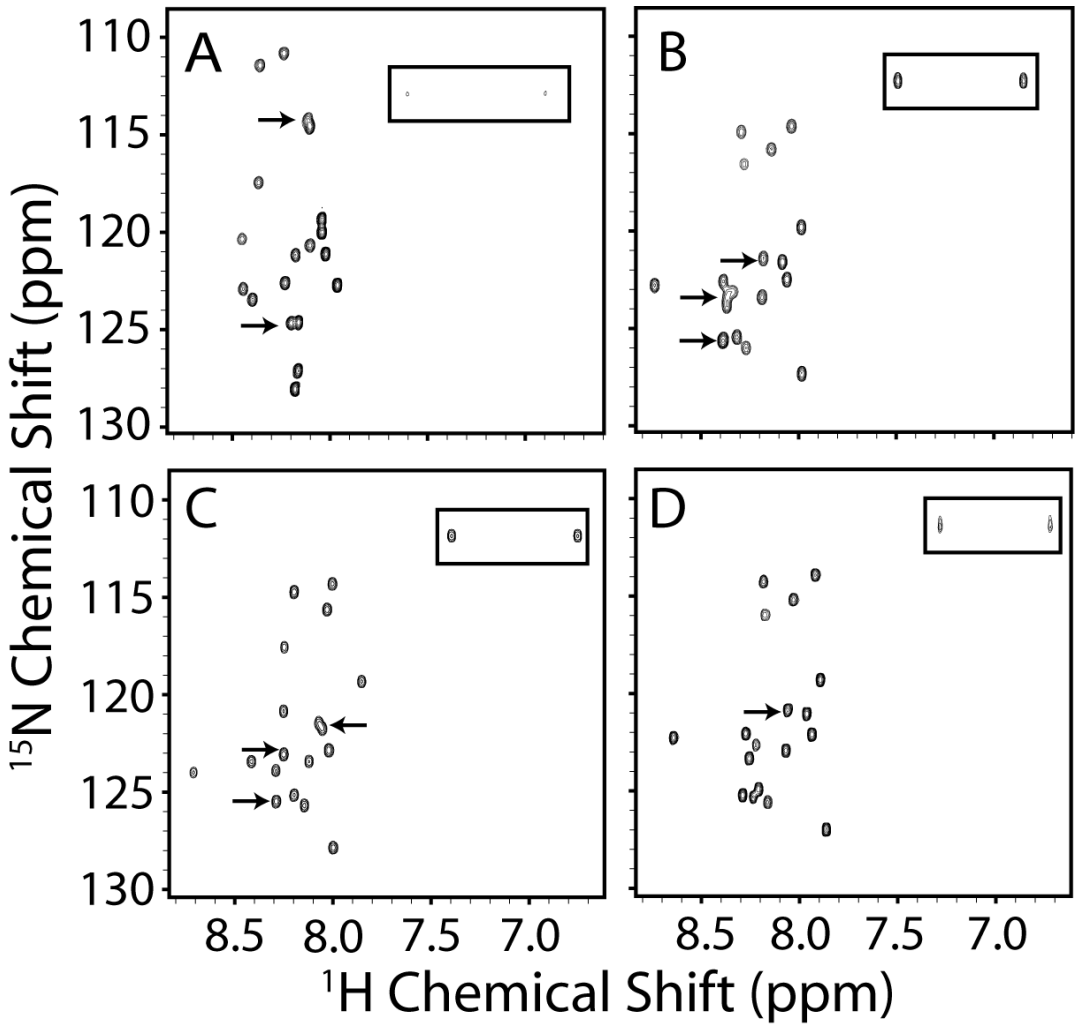


Figure 5: 1H/15N HSQC NMR spectra of 15N labeled peptides. A) PKI_{5-24}, B) PLN_{1-20}, C) R9C-PLN_{1-20}, and D) R14del-PLN_{1-20} at 1 mM total peptide concentration. The conditions used to obtain these spectra reflect the buffer and temperature in which PKAc has been previously studied (19). Arrows indicate unresolved peaks which were detected. Side chain amide peaks are indicated in the boxed regions.

and two mutants of PLN. Since our method uses *E. coli* as the host cell for recombinant expression, it permits all of the widely utilized and cost-efficient isotope labeling approaches in this host system for NMR spectroscopic analysis (20). By using MBP-fusion constructs, we were also able to limit the amount of proteolytic degradation of these peptides in *E. coli*. From one liter of media, our approach results in the production of 3 mg of PKI_{5-24} (0.5 mg peptide per gram wet cell mass), 2 mg of PLN_{1-20} or R14del-PLN_{1-20} (0.3 mg peptide per gram wet cell mass), and 1 mg R9C-PLN_{1-20} (0.2 mg peptide per gram wet cell mass). Thus, quantities sufficient for NMR structural and dynamic studies can be obtained in as little as one liter of media.

References

(1) Graves, J.D., Krebs, E.G. (1999) *Pharmacol. Ther.* 82, 111-121.

(2) Taylor, S.S., Knighton, D.R., Zheng, J., Ten Eyck, L.F., Sowadski, J.M. (1992) *Annu. Rev. Cell. Biol.* 8, 429-462.

(3) Johnson, D.A., Akamine, P., Radzio-Andzelm, E., Madhusudan, M., Taylor, S.S. (2001) *Chem. Rev.* 101, 2243-2270.

(4) Hauer, J.A., Barthe, P., Taylor, S.S., Parello, J., Padilla, A. (1999) *Protein Sci.* 8, 545-553.

(5) MacLennan, D.H., Kranias, E.G. (2003) *Nat. Rev. Mol. Cell. Biol.*4, 566-77.

(6) Zamoon, J., Mascioni, A., Thomas, D.D., Veglia, G., (2003) *Biophys. J.* 85, 2589-2598.

(7) Traaseth, N.J., Ha, K.N., Verardi, R., Shi, L., Buffy, J.J., Masterson, L.R., Veglia, G. (2008) *Biochemistry* 47, 3-13.

(8) Simmerman, H.K., Jones, L.R. (1998) *Physiol. Rev.* 78, 921-47.

(9) Schmitt,J.P., Kamisago, M., Asahi, M., Li, G.H., Ahmad, F., Mende, U., Kranias, E.G., MacLennan, D.H., Seidman, J.G., Seidman, C.E. (2003)*Science* 299, 1410-3.

(10) Haghighi, K., Kolokathis, F., Gramolini, A.O., Waggoner, J.R., Pater, L., Lynch, R.A., Fan, G.C., Tsiapras, D., Parekh, R.R., Dorn, G.W., MacLennan, D.H., Kremastinos, D.T., Kranias, E.G. (2006) *Proc. Natl. Acad. Sci.* 103, 1388-1393.

(11) Opella, S.J., Ma, C., Marassi, F.M. (2001) *Methods Enzymol.* 339, 285-313.

(12) Buck, B., Zamoon, J., Kirby, T.L., DeSilva, T.M., Karim, C., Thomas, D.D., Veglia, G. (2003) *Protein Expr. Purif.* 30, 253-261.

(13) Hu, J., Qin, H., Li, C., Sharma, M., Cross, T.A., Gao, F.P. (2007) *Protein Sci.* 16, 2153-2165.

(14) Dolja,V.V., Peremyslov, V.V., Keller, K.E., Martin, R.R., Hong, J. (1998) *Virology* 252, 269-274.

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

(16) Goddard, T.D., Kneller, D.G. (2006) SPARKY version 3.113, University of California, San Francisco.

(17) De Furia, F.G., Miller, D.R., Cerami, A., Manning, J.M. (1972) *J. Clin. Invest.* 51, 566-574.

(18) Material Safety Data Sheet Triton X-100 (2005) MSDS, MSDS No. SLT2452, Science Lab, Houston, Texas.

(19) Masterson, L.R., Mascioni, A., Traaseth, N.J., Taylor, S.S., Veglia, G. (2008) *Proc. Natl. Acad. Sci.* 105, 506-511.

(20) Lian, L.Y., Middleton, D.A., (2001) *Prog. NMR Spect.* 39, 171-267.

3. Carbonyl Carbon Label Selective (CCLS) ^{1}H-^{15}N HSQC Experiment for Improved Detection of Backbone ^{13}C-^{15}N Cross Peaks in Larger Proteins

Marco Tonelli, Larry R. Masterson, Klaas Hallenga, Gianluigi Veglia, and John L. Markley

Reprinted with permission from the Journal of Biomolecular NMR (2007)*39, 177-185.*

We present a highly sensitive pulse sequence, carbonyl carbon label selective $^{1}H–^{15}N$ HSQC (CCLS-HSQC) for the detection of signals from $^{1}H–^{15}N$ units involved in $^{13}C'–^{15}N$ linkages. The CCLS-HSQC pulse sequence utilizes a modified ^{15}N CT evolution period equal to $1/(2^{1}J_{NC'})$ (~33 ms) to select for $^{13}C'–^{15}N$ pairs. By collecting CCLS-HSQC and HNCO data for two proteins (8 kDa ubiquitin and 20 kDa HscB) at various temperatures (5 to 40 °C) to vary correlation times, our results demonstrated the superiority of the CCLS-HSQC pulse sequence for proteins with long correlation times (simulating higher molecular weight). We have then shown that the CCLS-HSQC experiment yielded assignments in the case of a 41 kD protein incorporating pairs of ^{15}N- and $^{13}C'$-labeled amino acids where a TROSY 2D-HN(CO) had failed. Although the approach requires that the ^{1}H-^{15}N HSQC cross peaks be observable, it does not require deuteration of the protein. The method is suitable for larger proteins and is less affected by conformational exchange than HNCO experiments, which require a longer period of transverse ^{15}N magnetization. The method also is tolerant to the partial loss of signal from isotopic dilution (scrambling). This approach will be applicable to families of proteins that have been resistant to NMR structural and dynamic analysis, such as large enzymes, and partially folded or unfolded proteins.

Introduction

Recent advances have enabled NMR investigations of proteins as large as 800 kDa (1). However, the *conditio sine qua non* for NMR structural, dynamic, and binding studies is the resonance assignment. Sequential assignment strategies rely on multidimensional experiments governed by magnetization transfer through scalar couplings that link spin systems (2). These methods can work for proteins labeled uniformly with ^{15}N and ^{13}C up to about 35 kDa. Larger proteins generally require the use of TROSY (transverse relaxation optimized spectroscopy) based pulse sequences (3) and perdeuteration combined with selective labeling (4, 5). However, these methods fall short when conformational exchange (6) or isotopic dilution effects (7) lead to sensitivity losses and incomplete spectral information (8-10). A useful approach with larger proteins is to incorporate pairs of ^{15}N- and ^{13}C′-labeled amino acids and to use one-bond ^{15}N-^{13}C′ correlations to assign linked dipeptides (11-14). However, unfavorable dynamics or dilution of the isotopic labels may make these correlations difficult to resolve by conventional NMR methods, such as a 2D-[^{1}H/^{15}N]-TROSY-HN(CO).

We propose here a new method for detecting ^{1}H-^{15}N signals associated with ^{15}N-^{13}C′ backbone pairs. The approach is suitable for use with a uniformly labeled protein, but is particularly useful with a selectively labeled larger protein, where it has supported assignments in a case where conventional methods failed. The method utilizes the carbonyl carbon label selective (CCLS) ^{1}H-^{15}N HSQC pulse sequence shown in Figure

1. It relies on a shorter magnetization transfer period between ^{15}N and $^{13}C'$ and is based on the detection of ^{1}H-^{15}N correlations without $^{13}C'$ evolution.

Materials and Methods

NMR Samples and Spectroscopic Conditions

This study made use of samples from three different proteins. The first protein we used was a uniformly doubly labeled sample of ubiquitin purchased from Martek Biosciences. The second protein was HscB from *Escherichia coli*, a 20 kDa J-type co-chaperone protein involved in iron sulfur cluster assembly. A sample of [U-^{13}C, U-^{15}N]-HscB produced from *E. coli* was generously provided by Dennis T. Ta and Larry E. Vickery (University of California, Irvine). The third protein was the 41-kDa C-subunit of mouse cAMP-dependent protein kinase A (PKA-C). We produced samples of PKA-C from *E. coli* cells, with two selective labeling patterns. Sample **A** was prepared by incorporating [$^{13}C'$]-Phe and [^{15}N]-Gly, and sample **B** was prepared by incorporating [$^{13}C'$]-Gly and [^{15}N]-Phe. PKA-C contains two Phe-Gly dipeptides (F54-G55, F185-G186), which should contain $^{13}C'$–^{15}N linkages in sample **A**, and one Gly-Phe dipeptide (G186-F187), which should contain a $^{13}C'$–^{15}N linkage in sample **B**.

NMR data were recorded on Varian Inova spectrometers equipped with a *z*-axis pulsed field gradient cold probe operating at ^{1}H Larmor frequencies of either 800 MHz (ubiquitin and PKA-C) or 600 MHz (HscB). Data for ubiquitin (1.5 mM, pH 6.5 in 50 mM phosphate buffer containing 7% D_2O) were acquired at 25 and 5 °C, while data for HscB (1 mM, pH 7.4 in 20 mM Tris buffer containing 7% D_2O and 9 mM DTT) were

acquired at 40, 30, 20, and 10 °C, and data for PKA-C (0.3 mM, pH 6.5 in 20 mM phosphate buffer containing 10% D_2O, 180 mM KCl and 15 mM DTT) were acquired at 27 °C.

2D CCLS-HSQC spectra were acquired on all protein samples using the pulse sequence shown in Figure 1 and compared with spectra acquired with a ^{1}H-^{15}N 2D-version of a conventional HNCO sequence for ubiquitin and HscB (15, 16), and a TROSY HNCO experiment for PKA-C (17). For the uniformly ^{15}N,^{13}C-labeled samples of HscB and ubiquitin, the CCLS-HSQC sequence shown in Figure 1 was modified as described in the figure legend to refocus $^{1}J_{NC\alpha}$ coupling during the ^{15}N constant-time evolution. For ubiquitin, all 2D CCLS-HSQC spectra were acquired with 4 scans and 64 complex points in the ^{15}N dimension, whereas 8 transients were used for each FID of the 2D HN(CO) spectra. For HscB, all 2D data were acquired with 16 transients and 53 complex points in the ^{15}N dimension, except for spectra collected at the lowest temperature (10 °C) for which 64 transients were accumulated for each FID. For the PKA-C samples, each 2D spectrum was collected using 640 scans and 40 complex points in the ^{15}N dimension. The software nmrPipe/nmrDraw (18) was used for data processing and initial visualization. Processing was achieved by zero filling to a final matrix size of 2048 x 128 points and applying a squared sine-bell window function shifted by 90° in both dimensions prior to Fourier transformation. Sparky software (19) was used for subsequent data visualization and for determining signal-to-noise estimates.

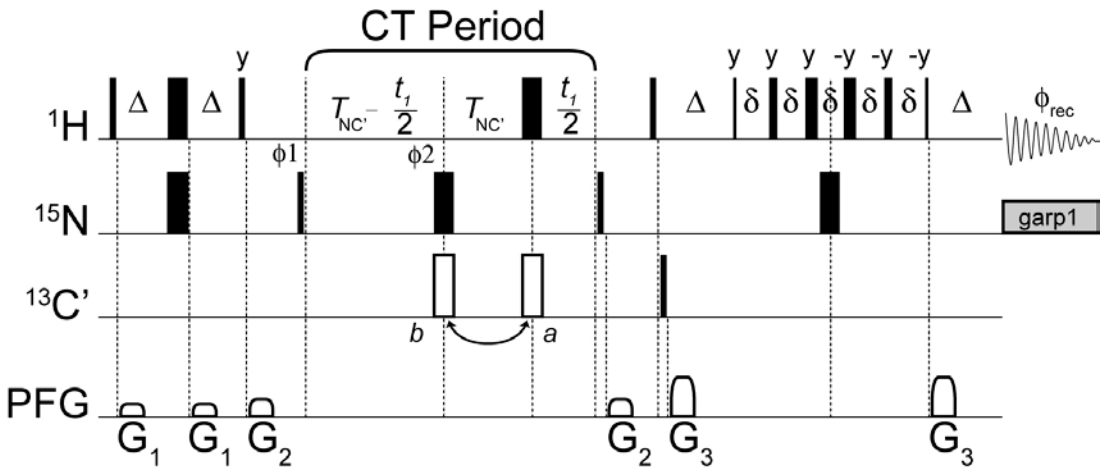


Figure 1: Schematic of the CCLS-HSQC pulse sequence. The reference spectrum is obtained by executing the pulse sequence with the 180° ^{13}C′ pulse (open rectangle) at position *a*; the ^{13}C′ suppressed spectrum is obtained with this pulse at position *b*. Unless indicated otherwise, all rectangular pulses are applied along the *x*-axis with a flip angle of 90° (narrow bars) or 180° (wide bars). The carrier frequency for ^{1}H is set on resonance with water at 4.77 ppm; the carrier frequency for ^{15}N is set in the center of the amide region at 121.8 ppm; for ^{13}C pulses the offset is set to the C′ region at 174.8 ppm. A 3-9-19 watergate pulse scheme is used in the reverse INEPT transfer to suppress the strong water signal. Garp1 decoupling is used on ^{15}N during acquisition. Delay durations: Δ = 2.4ms; δ = 0.11ms; $T_{NC'}$ = 16.5 ms. Phase cycling: ϕ_1 = *x*,-*x*, ϕ_2 = *x*,*x*,-*x*,-*x*, ϕ_{rec} = *x*,-*x*. To accomplish States quadrature detection for the ^{15}N indirect dimension, a second FID is collected for each increment by changing the ϕ_1 phase to *y*,-*y*. For States-TPPI acquisition, the ϕ_1 and ϕ_{rec} phases are also incremented by 180° every other ^{15}N increment. Gradients are Wurst shaped *z*-axis gradients with a length of 1 ms. Gradient strengths (G/cm): G1: 5, G2: 7, G3: 17. For the uniformly ^{15}N,^{13}C-labeled

HscB sample, in order to refocus $^{1}J_{NC\alpha}$ coupling during the ^{15}N constant-time evolution, the CCLS-HSQC pulse sequence was modified as follows: for the 90° and 180° pulses on $^{13}C'$, selective sinc shaped pulses were used with an excitation maximum at $^{13}C'$ and a null at $^{13}C^{\alpha}$, 113 ppm away; a selective squared 180° pulse with a maximum of excitation on $^{13}C^{\alpha}$ was given at point *a* to decouple ^{15}N and $^{13}C^{\alpha}$ nuclei during the constant-time evolution. In order to preserve water magnetization, our CCLS-HSQC pulse sequence is based on the fast HSQC experiment (20). According to this approach, water magnetization, which is placed on the transverse plane at the end of the first INEPT period, is dephased and rephased by the identical G2 gradients flanking the ^{15}N constant-time evolution period, and subsequently flipped back and preserved along the +Z axis for the reverse INEPT and acquisition periods. This is a very simple and robust experiment that works well for proteins.

The transverse relaxation rate, T_2, of the ^{15}N antiphase magnetization for ubiquitin and HscB was measured at various temperatures (*vide supra*) by recording a series of CCLS-HSQC reference spectra with increasing constant-time delays, $2T_{NC'}$, and fitting an exponential equation to the observed decaying signal. For HscB we collected 1D spectra, each accumulated with 1024 transients, and measured peak intensities by integrating over a region of the 1D ^{1}H spectra between 8.0 ppm and 9.4 ppm. For ubiquitin, instead, full 2D spectra were collected at various $T_{NC'}$ delays, allowing us to calculate the relaxation rate for each residue in the protein.

Description of the CCLS-HSQC pulse sequence

The CCLS-HSQC pulse sequence builds on the constant time (CT) HSQC experiment. Two spectra, a reference spectrum and a suppression spectrum, are acquired simultaneously in an interleaved manner. The reference spectrum is recorded using the pulse sequence shown in Figure 1, with the 180° pulse on ^{13}C′ during the ^{15}N CT evolution period (open rectangle) applied at position *a*, simultaneous to the 180° ^{1}H pulse. This is essentially a constant time HSQC experiment that yields a spectrum with decoupled ^{1}H-^{15}N signals. The suppression experiment is acquired with the 180° ^{13}C′ pulse at position *b*, simultaneous to the 180° ^{15}N pulse. This causes the ^{15}N-^{13}C′ coupling to be active during the constant time period so as to convert the transverse magnetization of ^{15}N spins attached to ^{13}C′ ($2H_zN_x$ and $2H_zN_y$) into antiphase magnetization, $4H_zN_yC'_z$ and $4H_zN_xC'_z$, respectively. Neither of these terms will yield observable signal at the end of the sequence. The $4H_zN_xC'_z$ term, which is not affected by the ^{15}N 90° pulse given along the *x*-axis at the end of the ^{15}N evolution, is dephased

by the G_2 gradient. The second term ($4H_zN_yC'_z$) is placed along the *z*-axis by the ^{15}N 90° pulse ($4H_zN_zC'_z$) and is not dephased by the G_2 gradient, but it is converted into unobservable multiple quantum coherence ($2H_xC'_y$) by the subsequent 90° pulses on ^{1}H and ^{13}C′ during the reverse INEPT period. As a consequence, signals from ^{1}H-^{15}N groups linked to ^{13}C′ are suppressed by this pulse scheme. On the other hand, magnetization from ^{12}C′-linked ^{1}H-^{15}N groups is unaffected by the suppression pulses on ^{13}C′ and results in cross-peaks that have the same intensity as the reference spectrum. This spectrum could then be subtracted from the reference one, to cancel out ^{12}C′-linked resonances, leaving only the ^{13}C′-linked ^{1}H–^{15}N peaks. However, as explained in more detail below, this causes a loss in the signal-to-noise ratio, S/N, for the CCLS method. An alternative and more advantageous strategy is not to subtract the data to obtain an edited spectrum, but rather to use the suppression spectrum only to identify which peaks in the reference spectrum are linked to ^{13}C′ nuclei.

A similar method was described earlier (21) in which the difference in intensity between a reference and a suppression spectrum is used for accurately extracting the three-bond ^{15}N-^{13}C *J* couplings. Although technically similar, the intended purpose of the two experiments is quite different.

Signal-to-noise comparison between CCLS-HSQC and HNCO experiments

Signal from the CCLS-HSQC reference spectrum is regulated only by the exponential decay of the transverse ^{15}N antiphase magnetization during the constant-time period ($2T_{NC'}$), $\exp(-2T_{NC'}/T_2)$. In the CCLS-HSQC suppression experiment, on the

other hand, the ^{15}N magnetization during the constant-time period is attenuated by the antiphase ^{15}N T_2 decay, $\exp(-2T_{NC'}/T_2)$, and modulated by the active $^1J_{NC'}$ coupling, $\cos(2\pi\ ^1J_{NC'}T_{NC'})$. This additional cosine modulation will result in complete suppression of the ^{13}C′-linked ^{1}H–^{15}N peaks for the suppression spectrum when the constant-time delay is equal to $1/(2\ ^1J_{NC'})$, or $T_{NC'} = 1/(4\ ^1J_{NC'})$.

If we subtract the suppression spectrum from the reference spectrum, the signal observed for the resulting edited spectrum is then given by the equation $[1-\cos(2\pi\ ^1J_{NC'}T_{NC'})]\bullet\exp(-2T_{NC'}/T_2)$, which can also be written as $2\bullet\sin^2(\pi\ ^1J_{NC'}T_{NC'})\bullet\exp(-2T_{NC'}/T_2)$. It can be easily shown from this formula that maximum signal will be obtained when the $T_{NC'}$ delay is set to $(1/\pi\ ^1J_{NC'})\bullet\text{atan}(\pi\ ^1J_{NC'}T_2)$. Surprisingly, this indicates that the maximum signal for the edited spectrum is obtained for constant-time periods that are longer than the $1/(2\ ^1J_{NC'})$ delay (or $T_{NC'} = 1/(4\ ^1J_{NC'})$).

In a similar fashion, the signal resulting from an HNCO experiment is governed by the equation $\sin^2(2\pi\ ^1J_{NC'}T_{NC'})\bullet\exp(-4T_{NC'}/T_2)$, where $T_{NC'}$ is the time delay for half ^{15}N-^{13}C′ INEPT transfer and is repeated four times during the HNCO pulse sequence. To maximize signal for HNCO experiments, the $T_{NC'}$ delay will have to be set to $(1/2\pi\ ^1J_{NC'})\bullet\text{atan}(\pi\ ^1J_{NC'}T_2)$, i.e. at shorter values than for the edited CCLS-HSQC spectrum above.

In order to compare the S/N of the CCLS-HSQC and HNCO experiments, one must account for the fact that the CCLS-HSQC method requires the collection of two spectra; thus for the same data collection time, the number of scans accumulated in the HNCO experiment will be twice the number of scans used in the reference and

suppression CCLS-HSQC data accumulation. Furthermore, when the suppression CCLS-HSQC spectrum is subtracted from the reference, the noise in the resulting edited spectrum adds up to the same level as for the HNCO that was collected using twice as many scans. Hence, for a correct comparison of the CCLS-HSQC reference, subtraction and edited spectra with the HNCO, the formula describing their respective sensitivities must be multiplied by factors $\sqrt{2}$, $\sqrt{2}$, 1 and 2, respectively. The relative S/N c urves for the HNCO and the CCLS-HSQC reference and edited spectra as a function of $T_{NC'}$ are shown in Figure 2 for different values of T_2. It is clear that, when $T_{NC'}$ is optimized for each experiment to achieve maximum S/N, the CCLS-HSQC reference spectrum is the most sensitive and more so for shorter $T_{NC'}$ values. This is even more remarkable if we consider that the CCLS-HSQC reference experiment is run for only half as long as the HNCO. On the other hand, due the increase in the noise level caused by the editing process, no gain in sensitivity is found for the CCLS-HSQC edited experiment over the HNCO.

Thus, in order to take advantage of the CCLS-HSQC experiment's higher sensitivity, the best strategy is to use the suppression spectrum to identify which peaks in the reference spectrum belong to the ^{15}N-^{1}H groups attached to ^{13}C′ nuclei, without actually subtracting the two spectra. Sensitivity is optimized by setting the constant time delay to $1/(2\ {}^{1}J_{NC'})$, $T_{NC'} = 1/(4\ {}^{1}J_{NC'})$ (or possibly shorter for larger proteins with very short T_2) to achieve maximum reduction of the signals in the suppression spectrum, while retaining as much signal as possible in the reference spectrum. As described earlier, the signal for this spectrum is regulated only by the exponential decay of the transverse ^{15}N antiphase magnetization during the constant time delay: $\exp(-2T_{NC'}/T_2)$.

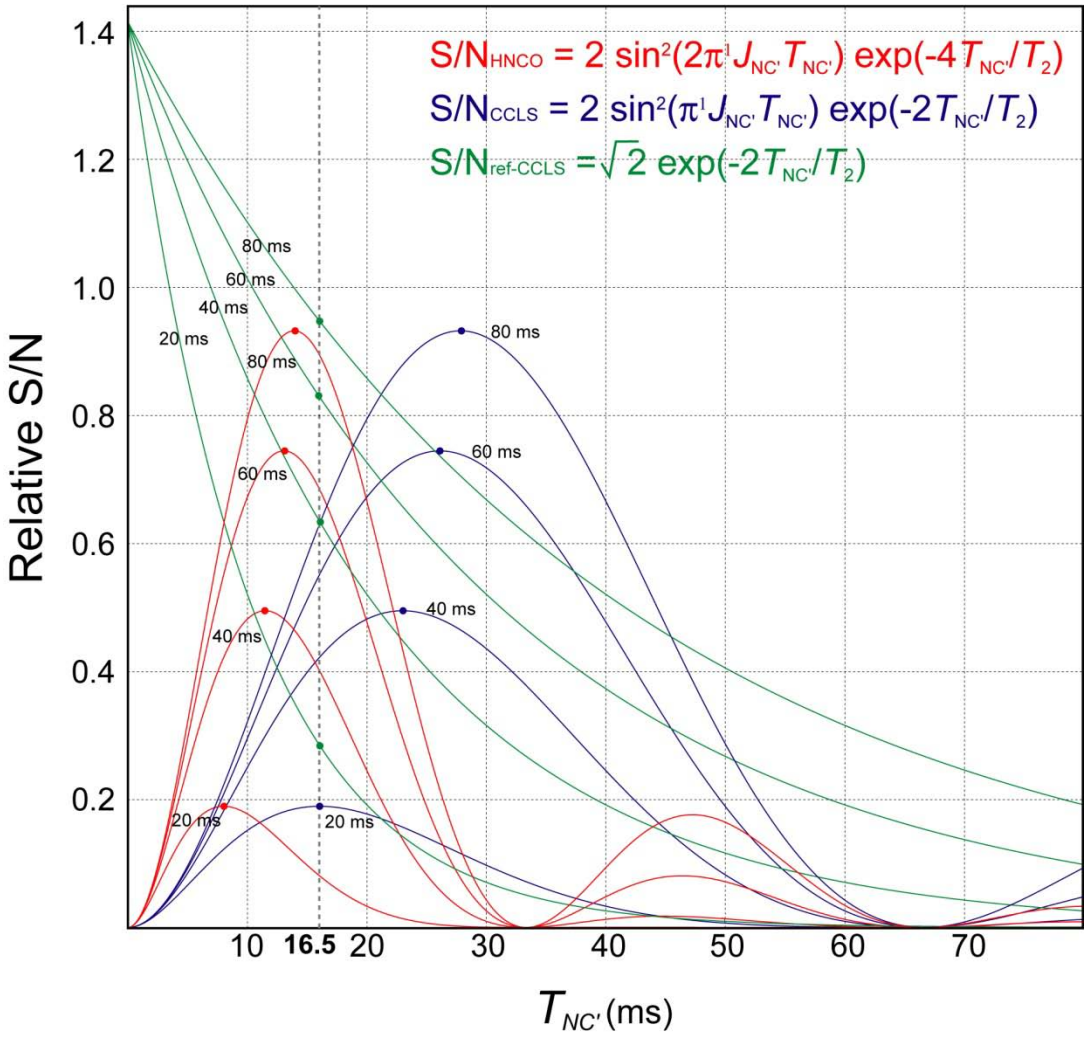


Figure 2: Plot showing the relative signal-to-noise ratio, S/N, for the edited CCLS-HSQC, reference CCLS-HSQC and HNCO spectra as a function of the $T_{NC'}$ delay and for four different values of the ^{15}N antiphase transverse relaxation time constant, T_2. The $T_{NC'}$ delay used in the equations corresponds to half of the constant-time delay for the CCLS-HSQC experiments and to half of the ^{15}N–^{13}C' INEPT transfers in the HNCO experiment. The colored dots on each curve indicate the maximum S/N obtainable for the edited CCLS-HSQC and HNCO experiments. For the reference

CCLS-HSQC experiment, instead, the dots are set at the $T_{NC'}$ value that results in complete suppression of the $^{1}H–^{15}N(–^{13}C')$ peaks in the suppression CCLS-HSQC spectrum. The equations that were used to calculate the curves are discussed in the Materials and Methods section and are also printed within the graph. This figure was generated using the commercially available program Grapher, version 1.0, from Apple Computer, Inc.

A careful analysis of the S/N curves for the various experiments (Figure 2) reveals that for large proteins, with shorter T_2, the S/N yielded by the reference spectrum is higher than that of either the HNCO or CCLS-HSQC edited spectrum acquired at the optimal $T_{NC'}$ value. The advantage of the CCLS-HSQC reference spectrum over the HNCO experiment for larger proteins becomes even more evident when we plot the ratio between the S/N of the two experiments (reference CCLS-HSQC over HNCO) as a function of the antiphase ^{15}N T_2 decay rate. This relationship, calculated using the theoretical formulas discussed above, is shown in Figure 3 along with our experimental results (see section below) and clearly demonstrates that the gain in sensitivity of the CCLS method over the HNCO steadily grows for proteins with ^{15}N T_2 values shorter than ~85 ms. Another advantage of the CCLS-HSQC spectra over the HNCO is that, for larger proteins, the optimal $T_{NC'}$ value that yields maximum signal for HNCO becomes very short allowing only very few increments to be collected within the constant-time ^{15}N evolution.

Results and Discussion

As an experimental verification of the theoretical equations describing the signal yielded by each experiment, we collected CCLS-HSQC reference and suppression spectra, as well as conventional HNCO spectra, for several different values of $T_{NC'}$, on a ubiquitin sample at two different temperatures: 5 and 25 °C. The S/N for each cross-peak in the 2D spectra was measured and plotted against the $T_{NC'}$ delay, and the theoretical curves described above were fitted to the experimental points of the corresponding spectrum, by optimizing two parameters: the ^{15}N transverse relaxation

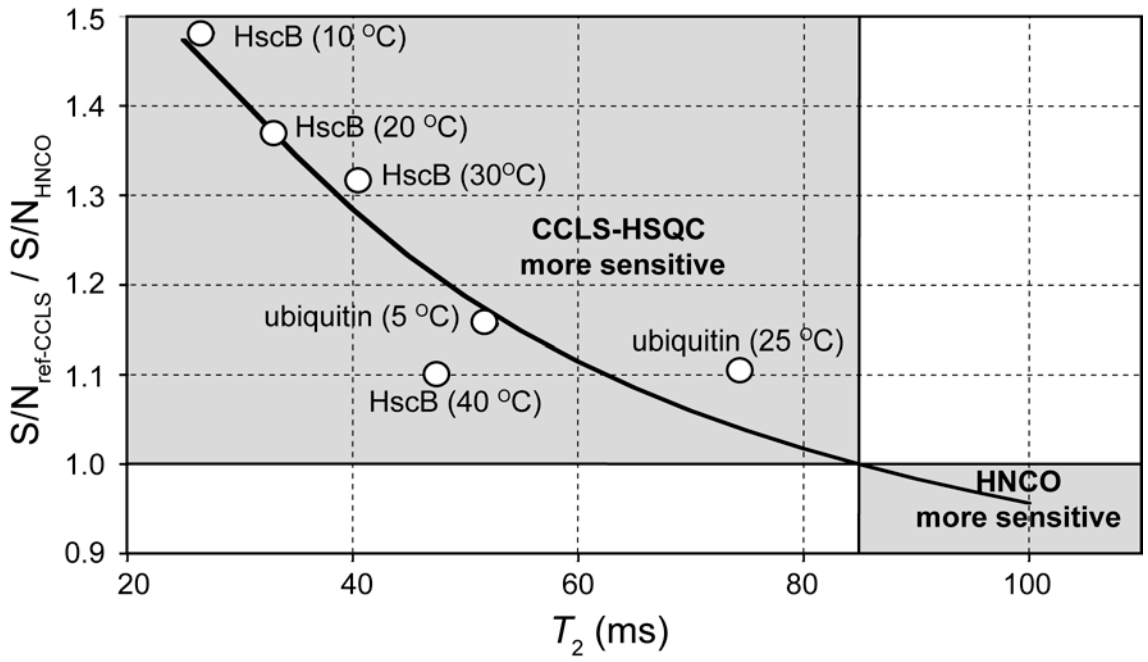


Figure 3: Plot showing how the sensitivity gain of the CCLS-HSQC reference spectrum over the HNCO spectrum (given by the ration between the S/N of the two experiments) changes as a function of the T_2 relaxation rate of the ^{15}N antiphase magnetization of the protein. The theoretical curve was calculated using the formulas described in the text using a $T_{NC'}$ delay of 16.5 ms for the CCLS-HSQC spectrum and using the optimized $T_{NC'}$ delay for any given T_2 value for the HNCO spectrum. The reported experimental points for ubiquitin were calculated by averaging the S/N ratios and the T_2 values calculated for each residue from 2D spectra. For HscB, the ^{15}N T_2 values were calculated by fitting the exponential decay of the signals comprised between 8.0 and 9.5ppm in a series of 1D CCLS-HSQC reference spectra acquired with increasing constant-time delays. The S/N ratios instead were calculated from 2D spectra as described in the text. This graph was made using KaleidaGraph.

time constant, T_2, and an arbitrary constant K. Figure 4 shows such a plot of measured S/N as a function of $T_{NC'}$ for a single residue, randomly chosen from the data set collected at 5 °C as representative of the quality of the experimental data, and the curve fits achieved. For this particular residue, curve fitting of the experimental points for all three experiments yielded very similar values of ^{15}N T_2 (~35 ms). It is also interesting to note that the predicted value of K for the different spectra (HNCO spectrum, CCLS-HSQC reference spectrum and CCLS-HSQC edited spectrum) reflects the same $2\sqrt{2}:1$ ratio described above. Similar results were obtained for the other ubiquitin residues at 5 and 25 °C.

As a test of the sensitivity of the CCLS method, we acquired reference CCLS-HSQC and conventional HNCO spectra on the 20-kDa [U-^{13}C, U-^{15}N]-HscB sample at 40, 30, 20, and 10 °C, corresponding to average T_2 values of ~47, 40, 33, and 27 ms. The signal-to-noise (S/N) ratio of each experiment decreased with lower temperature (Figure 5), as expected from the increase in rotational correlation time of the protein in solution. However, whereas the reference CCLS-HSQC sequence gave comparable signal to HNCO at 40 °C, the reference CCLS-HSQC experiment became more sensitive than the HNCO as the temperature was lowered to simulate a larger protein. At 10 °C, the conventional 2D-HN(CO) spectrum was quite poor, whereas the reference CCLS-HSQC spectrum still contained the majority of the cross peaks. At this temperature, the S/N ratios of reference CCLS-HSQC peaks were on average 50% higher than those of corresponding observable 2D-HN(CO) peaks (Table 1).

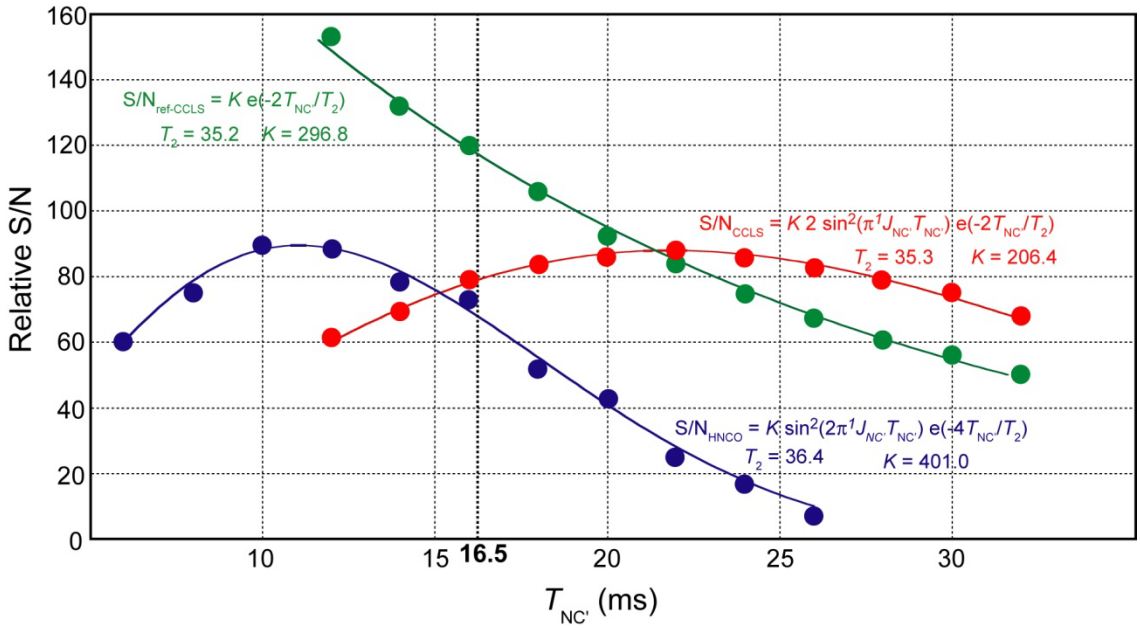


Figure 4: Plot showing the S/N for the edited CCLS-HSQC, reference CCLS-HSQC and HNCO spectra measured at several different values of the $T_{NC'}$ delay for a single residue of ubiquitin (randomly chosen). 2D-^{1}H/^{15}N spectra were collected, processed and analyzed as described in the Materials and Methods section. The HNCO spectra were accumulated using twice as many scans as the reference or suppression CCLS-HSQC spectra. The edited CCLS-HSQC spectra were obtained by subtracting the suppression spectra from the corresponding reference ones. The equations giving the S/N for each spectrum were fitted to the experimental points by optimizing T_2 and K. It is interesting to note that the predicted value of K for the different spectra reflects the same 2:√2:1 ratio between the HNCO, reference CCLS-HSQC spectrum and edited CCLS-HSQC spectrum, respectively, as described in the text. The theoretical equations as well as the optimized values of T_2 and K are also shown in the plot. For this particular residue, curve fitting of the data from the three spectra yielded very similar

values of ^{15}N T_2 (~35 ms). The ^{15}N T_2 value measured in this manner is the relaxation rate of the ^{15}N antiphase magnetization that is transverse during the constant-time period; it is this value of T_2 that plays a critical role in determining the signal intensities from the CCLS-HSQC and HNCO experiments. This plot and curve fitting were made using the commercially available program KaleidaGraph, version 3.6.2, by Synergy Software.

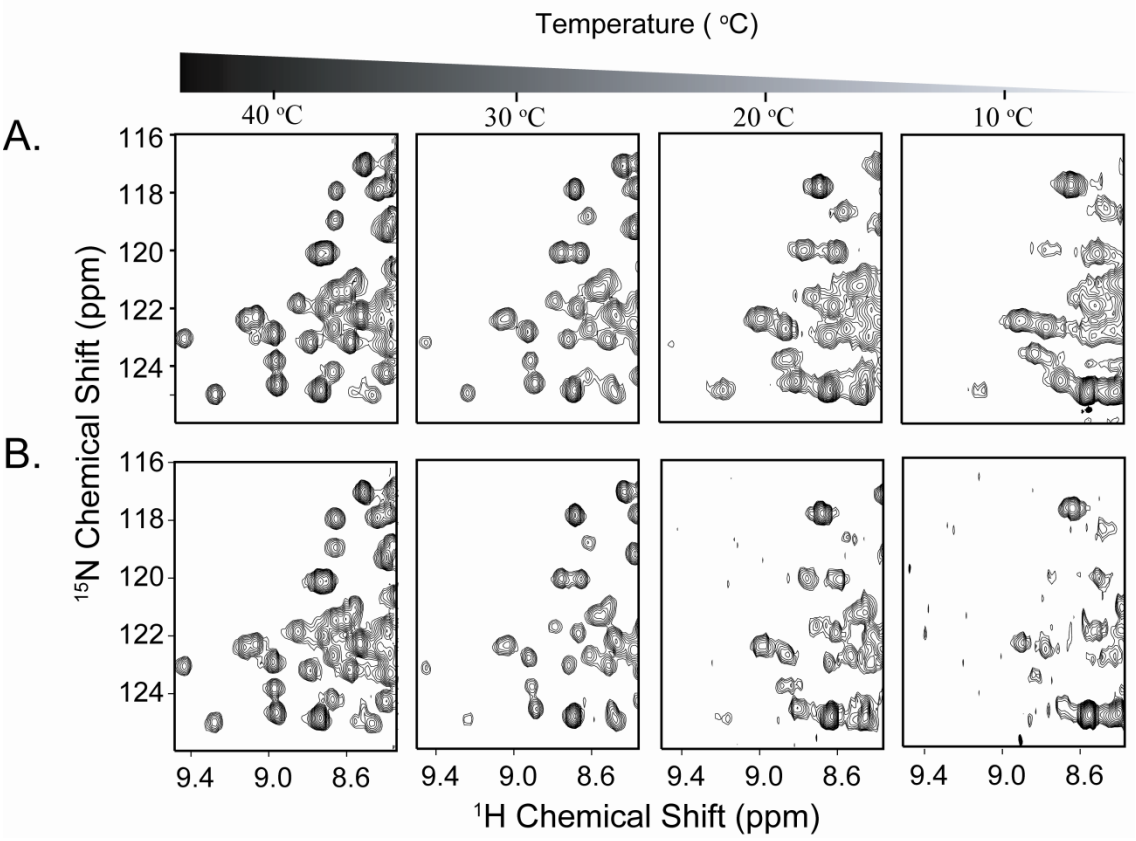


Figure 5: 2D-^{1}H/^{15}N spectra of the protein HscB taken at different temperatures using (A) the CCLS-HSQC pulse sequence and (B) the conventional 2D-HN(CO) pulse sequence. Identical acquisition parameters were used to collect the data using either sequence. Each FID was accumulated with 16 scans for all spectra, except for those run at 10 °C, where 64 scans were used. All processing was performed identically for each data set.

Table 1: NMR data for [U-^{13}C, U-^{15}N]-HscB acquired at various temperatures including the relative signal-to-noise ratios of cross peaks in the CCLS-HSQC and 2D-HN(CO) spectra.

Temperature (°C)	**40.0**	**30.0**	**20.0**	**10.0**
T_2 **(ms)**[a]	47.5 ± 0.1	40.5 ± 0.3	33.0 ± 0.3	26.6 ± 0.6
Relative S/N: CCLS-HSQC / 2D-HN(CO)[b]	1.10 ± 0.2	1.32 ± 0.3	1.37 ± 0.4	1.48 ± 0.5

[a]Relaxation times (T_2) were calculated as described in the Materials and Methods section.

[b]The signal-to-noise ratio was the average for the observed CCLS-HSQC cross peaks divided by the average for the corresponding 2D-HN(CO) cross peaks. The S/N of the CCLS-HSQC reference spectra was scaled down by $\sqrt{2}$ to account for the fact that the CCLS-HSQC and HNCO

The average of the S/N ratio between the spectra (CCLS-HSQC reference over HNCO) measured for HscB and ubiquitin at the various temperatures was then plotted (Figure 3) against the average of the corresponding antiphase ^{15}N T_2 rates. The experimental data points follow the same trend as the theoretical curve, calculated using the equations described earlier, clearly demonstrating that the CCLS approach is superior to the conventional 2D-HN(CO) experiment, with a significant sensitivity gain for large proteins.

We then applied the CCLS-HSQC method to the 41-kDa catalytic subunit of cAMP-dependent protein kinase A (PKA-C). PKA-C catalyzes the transfer of the γ-phosphate from ATP to a hydroxyl group of a serine or threonine residue. It is the most highly conserved member of the Ser/Thr kinase family and has therefore served as the prototype for this family of enzymes (22). Protein kinases are involved in a number of signal transduction cascades, including cell growth, proliferation, and death, making them ideal candidates as a target for drug design (23). This family of enzymes has remained relatively unexplored by NMR because of undesirable spectral characteristics related to the size of its members and the presence of conformational exchange effects on the μs–ms timescale (10, 24-26). Previous efforts to assign the resonances of PKA-C by TROSY-based 3D experiments yielded assignments to only ~55% of the non-proline residues (10). Among the missing assignments were highly conserved dynamic residues located in loops and turns in the active site of the enzyme directly involved in the catalytic cycle (22, 27). Specifically, residues within the glycine-rich loop and the DFG-motif central to enzyme function. Our own attempts to assign these residues (data not shown) either by 3D experiments on perdeuterated enzyme or by incorporating pairs of

^{15}N- and ^{13}C′-labeled amino acids with detection by 2D-[^{1}H/^{15}N]-TROSY-HN(CO) were unsuccessful.

The CCLS-HSQC reference spectrum for sample **A** (Figure 6A) revealed 29 resonances. Because the primary sequence of PKA-C contains 22 glycines, the additional signals arose from scrambling of the label from [^{15}N]-Gly. Biosynthetic interchange of labels between Gly and Ser is common in recombinant proteins produced from *E. coli* (7). Nevertheless, the suppression spectrum showed reduction in the intensity of resonances only from glycines linked to ^{13}C′-Phe. Only two peaks were identified by the suppression spectrum, one of which (G55) could be assigned based on 3D spectra; while the remaining resonance which does not have correlations in the 3D data sets was assigned to G186.

PKA-C contains only one Gly-Phe dipeptide, corresponding to G186 and F187. For sample **B**, the CCLS-HSQC reference spectrum (Fig. 6B**)** contained signals only from the 25 Phe residues in the PKA primary sequence, and hence showed no evidence for isotope scrambling. However, as demonstrated for sample **A**, the isotopic enrichment of ^{13}C′-Gly was diluted by biosynthetic exchange. Therefore, the F187 resonance intensity was reduced, rather than canceled, in the suppression CCLS-HSQC experiment. By comparing the reference and suppression CCLS-HSQC spectra, only one cross peak was identified which was then unambiguously assigned to F187. Assignments made in this manner have enabled us to map the chemical shift effects of nucleotide and substrate binding in the DFG-motif (28).

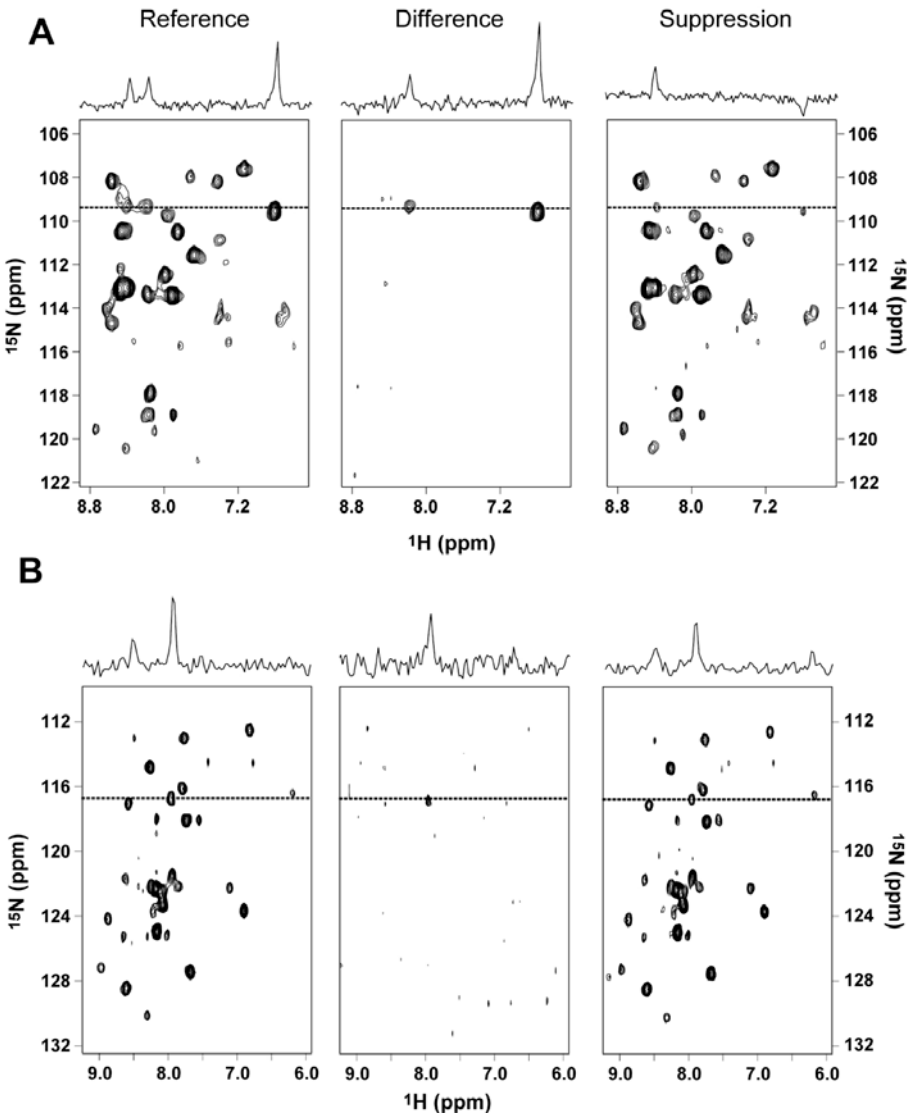


Figure 6: Spectra of the 41-kDa enzyme PKA-C utilizing the CCLS-HSQC sequence. (**A**) Sample **A** containing ^{15}N-Gly combined with ^{13}C′-Phe produced more than the expected resonances in the HSQC due to scrambling, but only 2 of the resonances expected peaks in the CCLS-HSQC. (**B**) Sample **B** containing ^{15}N-Phe combined with ^{13}C′-Gly produced 25 of the expected resonances in the HSQC spectrum and one of the expected resonances in CCLS-HSQC.

Conclusions

We present here a highly sensitive pulse sequence (CCLS-HSQC) for the detection of signals from $^{1}H–^{15}N$ units involved in a $^{13}C'–^{15}N$ linkage. A careful analysis of the formulas that govern the amount of signal obtained by the CCLS-HSQC and HNCO experiments indicates that the CCLS-HSQC pulse sequence provides higher sensitivity for larger proteins than the conventional HNCO. The difference becomes more obvious when the sensitivity gain provided by the CCLS-HSQC method over the conventional HNCO experiment is plotted as a function of the ^{15}N T_2 of the protein (Figure 3); the ratio of the S/N between the two spectra clearly favors the CCLS-HSQC experiment for ^{15}N T_2 values shorter than ~85 ms, with bigger gains achieved for larger proteins with shorter ^{15}N T_2.

As a verification of the sensitivity of the method, we acquired several 2D CCLS-HSQC and HNCO spectra at various temperatures on ubiquitin and a 20-kDa-protein sample. As shown in Figure 4, the experimental data we obtained is in good agreement with our theoretical considerations, further proving that the CCLS-HSQC experiment is more sensitive than HNCO for larger proteins. We then used the CCLS-HSQC experiment on selectively labeled PKA-C samples to assign a critical residue in a case where conventional 2D-HN(CO) had failed (28). Although the approach requires that the ^{1}H-^{15}N HSQC cross peaks be observable, it does not require deuteration of the protein. The method is suitable for larger proteins and is less affected by conformational exchange than HNCO experiments, which require a longer period of transverse ^{15}N magnetization. The method also is tolerant to the partial loss of signal from isotopic

dilution (scrambling). Finally, CCLS-HSQC can be used with combinatorial selective labeling schemes (29) and algorithms that provide optimized choices for potential assignments from $^{15}N/^{13}C'$ labeling (30). This approach will allow NMR investigations of families of proteins that have hitherto been resistant to NMR structural and dynamic analysis, such as large enzymes and unfolded proteins.

References

(1) Riek R, Fiaux J, Bertelsen EB, Horwich AL & Wuthrich K (2002) *J Am Chem Soc* 124, 12144-12153.

(2) Clore GM & Gronenborn AM (1994) *Meth Enzymol* 239, 349-363.

(3) Pervushin K, Riek R, Wider G & Wuthrich K (1997) *Proc Natl Acad Sci U S A* 94, 12366-71.

(4) Gardner KH & Kay LE (1998). *Annu Rev Biophys Biomol Struct* 27, 357-406.

(5) Tugarinov V, Hwang PM & Kay LE (2004) *Annu Rev Biochem* 73, 107-46.

(6) Wagner G (1993) *J Biomol NMR* 3, 375-385.

(7) Lian LY & Middleton DA (2001) *Progress in Nuclear Magnetic Resonance Spectroscopy* 39, 171-267.

(8) Abdulaev NG, *et al* (2006) *Biochemistry (N Y)* 45, 12986-12997.

(9) Tzakos AG, Grace CR, Lukavsky PJ & Riek R (2006) *Annu Rev Biophys Biomol Struct* 35, 319-342.

(10) Langer T, *et al* (2004) *Chembiochem* 5, 1508-1516.

(11) Kainosho M & Tsuji T (1982) *Biochemistry* 21, 6273-6279.

(12) McIntosh LP & Dahlquist FW (1990) *Q Rev Biophys* 23, 1-38.

(13) Tzakos AG, Grace CR, Lukavsky PJ & Riek R (2006) *Annu Rev Biophys Biomol Struct* 35, 319-342.

(14) Yabuki T, *et al* (1998) *J Biomol NMR* 11, 295-306.

(15) Kay LE, Ikura M, Tschudin R & Bax A (1990) *Journal of Magnetic Resonance* 89, 496-514.

(16) Ikura M, Kay LE & Bax A (1990) *Biochemistry* 29, 4659-4667.

(17) Weigelt J (1998) *J Am Chem Soc* 120, 10778-10779.

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

(19) Goddard TD & Kneller DG (2006) SPARKY3.113; University of California, San Francisco.

(20) Mori S, Abeygunawardana C, Johnson MO & van Zijl PC (1995) *J Magn Reson B* 108, 94-98.

(21) Vuister GW, Wang AC & Bax A (1993) *J Am Chem Soc* 115, 5334-5335.

(22) Taylor SS, *et al* (2005) *Biochim Biophys Acta* 1754, 25-37.

(23) Taylor SS, Knighton DR, Zheng J, Ten Eyck LF & Sowadski JM (1992) *Annu Rev Cell Biol* 8, 429-462.

(24) Langer T, *et al* (2005) *FEBS Lett* 579, 4049-4054.

(25) Vogtherr M, *et al* (2005) *J Biomol NMR* 32, 175.

(26) Vogtherr M, *et al* (2006) *Angew Chem Int Ed Engl* 45, 993-997.

(27) Taylor SS, *et al* (2004) *Biochim Biophys Acta* 1697, 259-269.

(28) Masterson LR, Mascioni A, Traaseth NJ, Taylor SS & Veglia G (2008) *Proc Natl Acad Sci U S A*

(29) Parker MJ, Aulton-Jones M, Hounslow AM & Craven CJ (2004) *J Am Chem Soc* 126, 5020-5021.

(30) Trbovic N, *et al* (2005) *J Am Chem Soc* 127, 13504-13505.

4. Simultaneous Detection and Deconvolution of Congested NMR Spectra Containing Three Isotopically Labeled Species

Larry R. Masterson, Marco Tonelli, John L. Markley, and Gianluigi Veglia

Reprinted with permission from the Journal of the American Chemical Society (2008) *130, 7818-7819.*

We present a procedure for isolating subspectra corresponding to individual protein or peptide components in a ternary mixture or complex. Each of the three-component species is labeled differently: species A uniformly with ^{15}N, species B uniformly with ^{15}N and ^{13}C, and species C uniformly with ^{15}N but selectively with $^{13}C'$ or $^{13}C^{\alpha}$. By using the dual carbon label selective HSQC (DCLS-HSQC) pulse sequence and exploiting differences in ^{1}J $^{15}N–^{13}C$ coupling patterns to filter selected ^{15}N resonances from detection during a constant time period, a subspectrum from each species can be generated from three spectra acquired from a single sample. Many important biological pathways involve dynamic interactions among members of multicomponent protein assemblies, and this approach offers a powerful way to monitor such processes.

Introduction

Analyses of protein fingerprint NMR spectra (e.g. 2D ^{1}H-^{15}N correlated spectra) yield atomic portraits of the conformational and dynamic changes that accompany binding events, such as ligand-protein, protein-protein, protein-DNA, and protein-RNA interactions (1-6). The introduction of NMR techniques such as TROSY (transverse relaxation optimized spectroscopy) (7) and CRINEPT (cross-relaxation enhanced polarization transfer) (8) has extended these approaches to biomolecules as large as 1 MDa (9). A common strategy in such experiments is to label only one binding partner with ^{15}N so that its signal is detected while that from the unlabeled partner remains 'NMR silent'. Then, a complex is prepared with another partner labeled so that its signals can be observed. However, it would be ideal to detect all species using *one experiment and one sample preparation*, so as to avoid the cost of preparing a second sample and to eliminate inconsistencies between sample preparations. Recently, Bermel *et al.* proposed such an approach for a two-component complex: one partner is ^{15}N-labeled and the other is double labeled with ^{13}C and ^{15}N, and a modified HNCO experiment is used to simultaneously detect and obtain sub-spectra of the two species (10). With a ternary complex (*e.g.*, enzyme-substrate-inhibitor), however, this approach would leave one species 'NMR silent'.

In this communication, we present an approach for the simultaneous detection and deconvolution of $^{1}H/^{15}N$ correlations in a ternary mixture of isotopically labeled proteins. This approach will support the molecular analysis of the myriad biological

systems that involve multiple species, *e.g.*, disruptions of protein-protein complexes by peptides; competitive, cooperative, and allosteric binding; and protein assemblies.

Materials and Methods

To separate the spectrum from each species in a ternary mixture, we used species-selective isotopic labeling in concert with a new pulse sequence, *D*ual *C*arbon *L*abel *S*elective (DCLS) $^{1}H/^{15}N$ HSQC (Figure 1A), based on spin-echo filtering (11, 12). The labeling scheme for the ternary mixture is as follows: (species A) uniformly ^{15}N labeled, (species B) uniformly ^{13}C and ^{15}N labeled, and (species C) ^{15}N,$^{13}C'$ labeled. The DCLS-$^{1}H/^{15}N$-HSQC enables the filtering of each species from detection by exploiting differences in ^{1}J scalar coupling of ^{15}N nuclei during a constant time (CT) period (Figure 1A). A series of three interleaved experiments are acquired in parallel. First, a reference data set is collected with simultaneous 180º selective pulses on $^{13}C'$ and $^{13}C^{\alpha}$ concurrent with a 180º square pulse on $^{1}H^{N}$ during the ^{15}N CT evolution (all pulses at positions "***a***"). This scheme results in suppression of frequency modulation arising from $^{1}J_{N\text{-}C'}$, $^{1}J_{N\text{-}C\alpha}$, and $^{2}J_{N\text{-}C\alpha}$ scalar couplings, and ^{15}N signals are thus decoupled from any bonded ^{13}C nucleus. In this case, all $^{1}H/^{15}N$ correlations are observed as in a conventional decoupled $^{1}H/^{15}N$ HSQC spectrum. Next, the first suppression data set is acquired by using a pulse sequence in which the 180º selective pulse on $^{13}C'$ is moved from position ***a*** to position ***b***. At the beginning of the CT delay (point I), ^{1}H-^{15}N magnetization is present as $2H_zN_{x/y}$ (depending on the phase of the preceding ^{15}N 90º pulse). During the CT delay, frequency modulation of ^{15}N magnetization due to active $^{1}J_{NC'}$ coupling occurs, and any ^{15}N resonance attached to

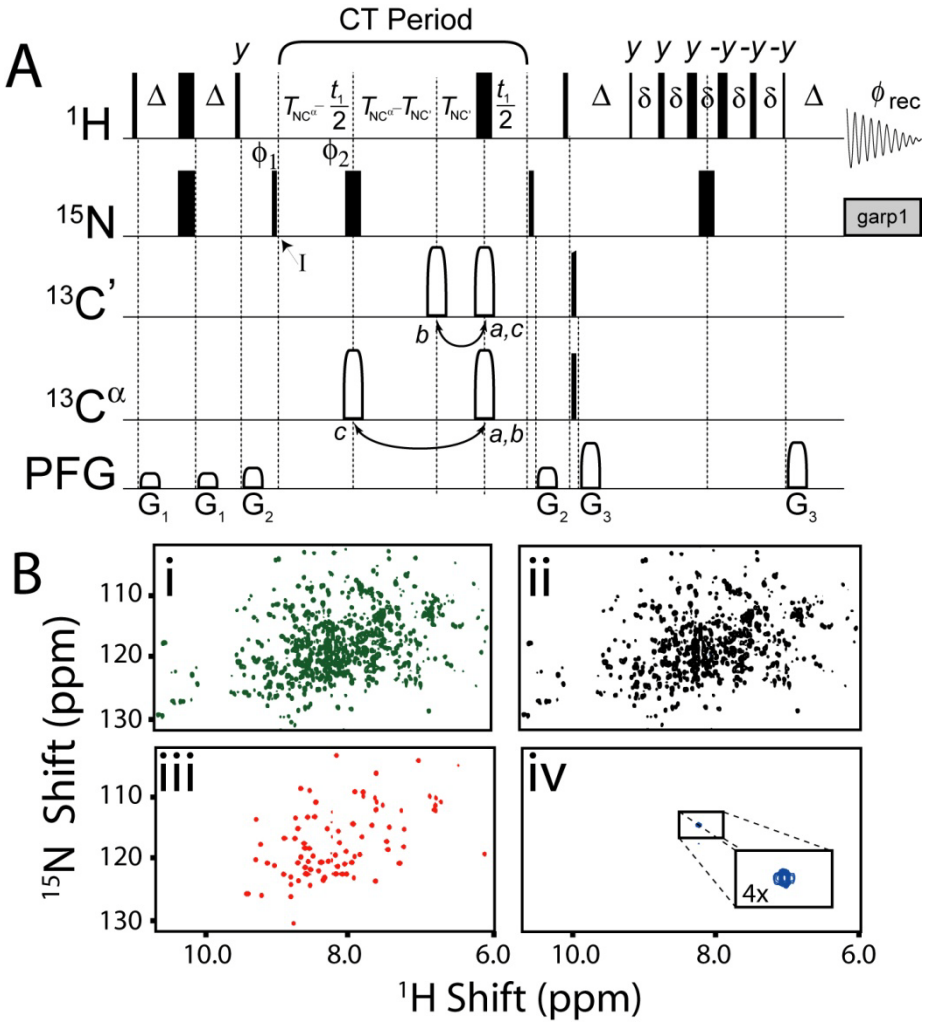


Figure 1:. A) The DCLS-HSQC pulse sequence. Optimal suppression of both $^{13}C^{\alpha}$ and $^{13}C'$ attached NH groups use $T_{NC\alpha}$ and $T_{NC'}$ delays of 24.5 and 16.4 ms, respectively (CT period of 49 ms). Further pulse sequence details in Supporting Information. **B**) Representative spectra obtained at 37 °C on a Varian VNMRS 800 MHz spectrometer equipped with a cryogenic probe: i. full ternary mixture consisting of 1 mm [U-^{2}H, U-^{15}N]MBP, 0.8 mM [U-^{13}C, U-^{15}N]-ubiquitin, and 0.8 mM [^{13}C′]-Ala4, [^{15}N]-Ser5 Kemptide; ii. MBP subspectrum; iii. ubiquitin subspectrum; iv. Kemptide subspectrum.

$^{13}C'$ evolves into antiphase magnetization ($4H_zN_{x/y}C'_z$). The residual $2H_zN_{x/y}$ term is modulated by $\cos(2\pi^1J_{NC'}T_{NC'})$ and can be completely suppressed by setting $T_{NC'}$ to $1/4^1J_{NC'}$, *i.e.* 16.4 ms (11). The antiphase terms are either dephased by gradient G_2 (for the $4H_zN_xC'_z$ term), or converted to unobservable multiple quantum coherence during the reverse INEPT transfer (for the $4H_zN_yC'_z$ term). This suppression scheme provides selective observation of signals from species A, whose NH groups are not attached to a $^{13}C'$. Finally, a second suppression experiment is acquired by moving the 180º selective pulse on $^{13}C^{\alpha}$ from position ***a*** to position ***c***. Under active $^1J_{NC\alpha}$ and $^2J_{NC\alpha}$ coupling, $2H_zN_{x/y}$ magnetization arising from $^{15}N,^{13}C^{\alpha}$ labeled species evolves into various antiphase magnetization terms ($4H_zN_{x/y}C^{\alpha(\imath)}_z$, $4H_zN_{x/y}C^{\alpha(i-1)}_z$, $8H_zN_{x/y}C^{\alpha(\imath)}_zC^{\alpha(i-1)}_z$) that are either dephased by the G_2 gradient or converted into unobservable multiple quantum coherences during the reverse INEPT transfer. Any observable signal arising from the residual $2H_zN_{x/y}$ term is modulated by the function $\cos(2\pi^1J_{NC\alpha}T_{NC\alpha})\bullet\cos(2\pi^2J_{NC\alpha}T_{NC\alpha})$. Since $^1J_{NC\alpha}$ is larger than $^2J_{NC\alpha}$, the $2H_zN_{x/y}$ term can be suppressed by setting $T_{NC\alpha}$ to $1/4^1J_{NC\alpha}$. However, whereas $^1J_{NC'}$ varies little with conformation, $^1J_{NC\alpha}$ and $^2J_{NC\alpha}$ show small dependence on protein backbone conformation (13, 14). Wirmer and Schwalbe (14) reported values of 10.8 ± 0.8 Hz for $^1J_{NC\alpha}$ and 7.7 ± 1.3 Hz for $^2J_{NC\alpha}$ for ubiquitin, with a positive correlation between the two coupling constants (i.e. larger values of $^1J_{NC\alpha}$ correspond to larger values of $^2J_{NC\alpha}$). Based on this, the suppression of signals for a $^{15}N,^{13}C^{\alpha}$ labeled species is best achieved by setting $T_{NC\alpha}$ to 24.5 ms. This completely suppresses signals from small values of $^1J_{NC\alpha}$, while signals with large values of $^1J_{NC\alpha}$

are inverted with residual intensities less than 4% (see Supporting Information). The only observable correlations will be those without $^{13}C^{\alpha}$ labeling (species A and C).

Results and Discussion

As a proof of concept, we applied this labeling scheme and pulse sequence to a ternary mixture of non-interacting proteins: maltose binding protein (MBP, 370 residues), ubiquitin (76 residues), and Kemptide (7 residues). Recombinant proteins were expressed in *Escherichia coli* BL21(DE3) to obtain MBP uniformly labeled with ^{2}H and ^{15}N and ubiquitin uniformly labeled with ^{13}C and ^{15}N. Kemptide was prepared by solid phase peptide synthesis using standard Fmoc chemistry. Labeling was introduced by incorporating $^{13}C'$ labeled Fmoc-alanine (at residue 4) and ^{15}N labeled Fmoc-serine(tBu) (at residue 5). The NMR sample contained 1 mM MBP, 5 mM Kemptide, and 0.8 mM ubiquitin.

Resulting spectra acquired with the DCLS-HSQC pulse sequence are in Figure 1B. All ^{1}H-^{15}N correlations were observed in the ^{13}C-decoupled DCLS-HSQC reference spectrum (Figure 1Bi**)**. As expected, the first suppression experiment resulted in the filtering of both $^{13}C'$ labeled species from detection, and only correlations attributed to MBP were observed **(**Figure 1Bii**)**. The second suppression experiment selectively filtered ubiquitin from detection. Separate sub-spectra for ubiquitin and Kemptide were obtained by linear combinations of the data sets: subtraction of the reference spectrum from the second suppression spectrum (^{15}N-$^{13}C^{\alpha}$ filtered detection) provided a sub-spectrum for ubiquitin (Figure 1Biii), and subtraction of the second suppression spectrum (^{15}N-$^{13}C^{\alpha}$ filtered) from the first suppression spectrum provided a sub-

spectrum for Kemptide (Figure 1Biv). In this manner, subspectra were obtained from a single sample for each individual component of the ternary mixture, and all of the resonances could be resolved for chemical shift analysis.

Introduction of selectively labeled $^{13}C'$ can be accomplished in recombinant proteins by using ^{15}N and $^{13}C'$ labeled amino acids, or in a more cost efficient manner, by using 1-^{13}C pyruvate and ^{13}C-labeled $NaHCO_3$ as the sole carbon sources (15). Alternatively, selective $^{13}C^{\alpha}$ labeling can be achieved by using 2-^{13}C glucose as the sole carbon source (16). In either case, a boost in sensitivity can be obtained using Perdeuteration (17, 18).

Our method offers advantages over the recently proposed HNCO-based approach (10). First, it enables the detection and deconvolution of spectra from three different species simultaneously. Second, the pulse sequence introduced here doesnot rely on complete refocusing of $^{13}C'$-^{15}N coupled nuclei for detection; thus the CT period in which ^{15}N magnetization is transverse is shorter, alleviating losses due to T_2 relaxation (49 vs 66 ms). If the HNCO-based approach were adapted to observe ^{15}N-$^{13}C^{\alpha}$ species, the CT period would be ~100 ms. In addition; the $^{13}C'$-filtered subspectrum for the ^{15}N labeled species (A) has a S/N ratio that is $\sqrt{2}$ higher than the corresponding HNCO-based spectrum. Due to spectral editing however, the subspectra for the two ^{13}C labeled species (B and C) have S/N ratios comparable to the HNCO-based approach for ^{15}N-$^{13}C'$ labeled species (10). Therefore, our recommendation is that the largest protein in the complex under investigation should only be ^{15}N labeled. Finally, additional labeling patterns should allow the possibility to resolve separate spectra from other species. A variety of selective labeling technologies, including

asymmetric methyl labeling (Ile vs. Val plus Leu) (19), segmental labeling (20), or cell-free labeling (21), could be combined with the DCLS-HSQC scheme to support NMR investigations of assemblies containing five or more species.

Conclusion

In summary, the highly sensitive pulse sequence and labeling scheme presented here supports the simultaneous detection and deconvolution of spectra from individual members of a ternary mixture of proteins. Many important biological pathways involve dynamic interactions among members of multi-component protein assemblies, and this approach offers a powerful way to monitor such processes.

References

(1) Betz, M.; Saxena, K.; Schwalbe, H. (2006) *Curr. Opin. Chem. Biol. 10,* 219-225.

(2) Kay, L. E. (2005) *J. Magn. Reson. 173,* 193-207.

(3) Clarkson, J.; Campbell, I. D. (2003) *Biochem. Soc. Trans. 31,* 1006-1009.

(4) Sitkoff, D.; Case, D. A. (1997) *J. Am. Chem. Soc. 119,* 12262-12273.

(5) Revington, M.; Holder, T. M.; Zuiderweg, E. R. (2004) *J. Biol. Chem. 279,* 33958- 33967.

(6) Matsuo, H.; Walters, K. J.; Teruya, K.; Tanaka, T.; Gassner, G. T.; Lippard, S. J.; Kyogoku, Y.; Wagner, G. (1999) *J. Am. Chem. Soc 121,* 9903-9904.

(7) Pervushin, K.; Riek, R.; Wider, G.; Wuthrich, K. (1997) *Proc. Natl. Acad. Sci. USA 94,* 12366-71.

(8) Riek, R.; Wider, G.; Pervushin, K.; Wuthrich, K. (1999) *Proc. Natl. Acad. Sci. USA 96,* 4918-4923.

(9) Sprangers, R.; Kay, L. E. (2007) *Nature 445,* 618-622.

(10) Golovanov, A. P.; Blankley, R. T.; Avis, J. M.; Bermel, W. (2007) *J. Am. Chem. Soc.129,* 6528-6535.

(11) Vuister, G.W.; Wang, A.C.; Bax, A.; (1993) *J. Am. Chem. Soc.* 115, 5334-35.

(12) Tonelli, M.; Masterson, L.R.; Hallenga, K; Veglia, G.; Markely, J.L. (2007) *J. Biomol. NMR* 39, 177-185.

(13) Delaglio, F.; Torchia, D.A.; Bax, A. (1991) *J. Biomol. NMR* 1, 439-446.

(14) Wirmer, J.; Schwalbe, H.; (2002) *J. Biomol. NMR* 23, 47-55.

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

(16) Lundstrom, P.; Teilum, K.; Carstensen, T.; Bezsonova, I.; Wiesner, S.; Hansen, D. F.; Religa, T. L.; Akke, M.; Kay, L. E. (2007) *J. Biomol. NMR 38*, 199-212.

(17) Fiaux, J.; Bertelsen, E. B.; Horwich, A. L.; Wuthrich, K. (1995) *J. Biomol. NMR 5*, 339-44.

(18) Pitto Venters, R. A.; Huang, C. C.; Farmer, B. T.,2nd; Trolard, R.; Spicer, D.; Fierke, C. A. (1995) *J Biomol NMR 5*, 339-44.

(19) Traaseth, N. J.; Verardi, R.; Veglia, G. (2008) *J. Am. Chem. Soc. 130*, 2400-2401.

(20) Xu, R.; Ayers, B; Cowburn, D.; Muir, T.W.; (1999) *Proc. Natl. Acad. Sci. USA* 96, 388-393.

(21) Yabuki, T.; Kigawa, T.; Dohmae, N.; Takio, Kl; Terada, T.; Ito, Y.; Laue, E.D.; Cooper, J.A.; Kainosho, M.; Yokoyama, S.; (1998) *J. Biomol. NMR* 11, 295-306.

Supporting Information

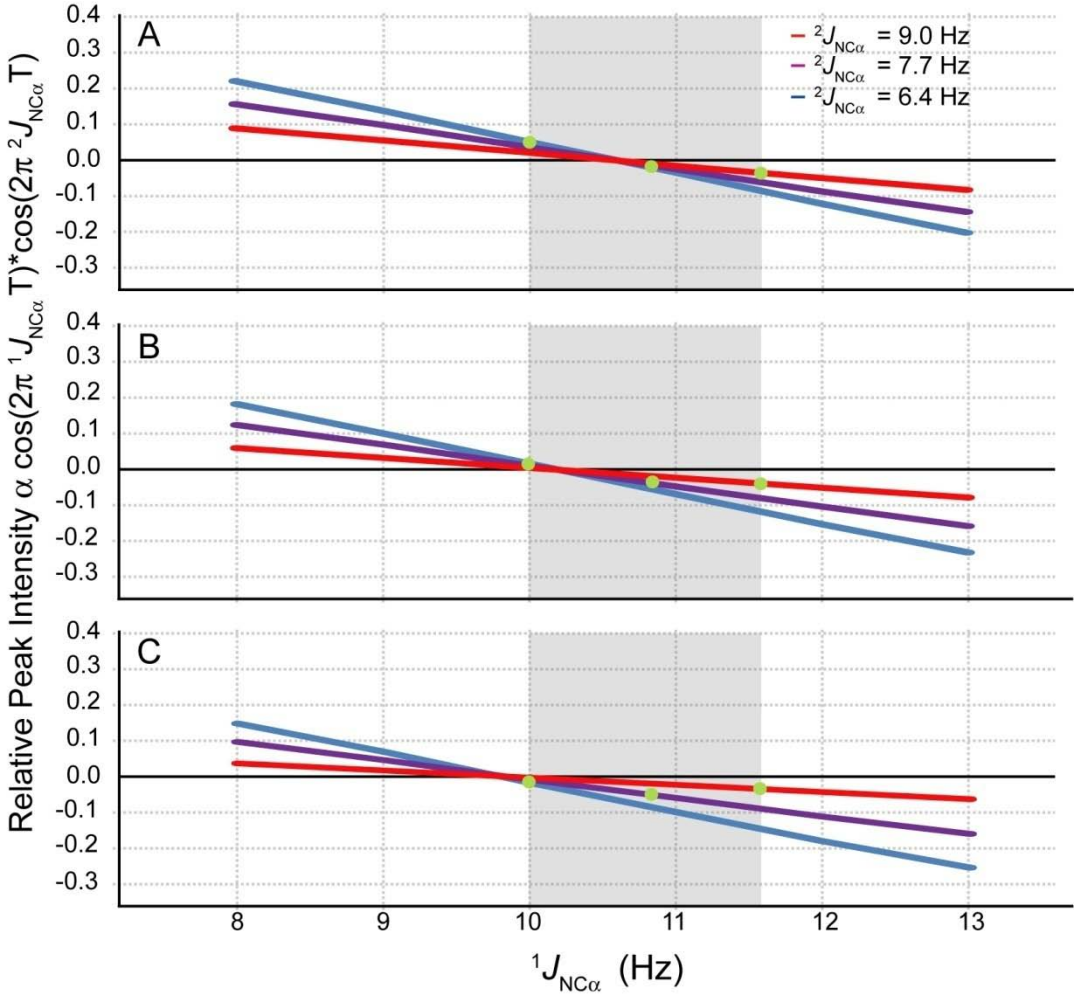


Figure S1: Optimization of the constant time period for the suppression of $^{13}C^{\alpha}$-attached ^{15}N resonances. For resonances that are involved in ^{15}N-$^{13}C^{\alpha}$ coupling, the relative peak intensity (neglecting T_2 relaxation) during the CT period is determined by the function $\cos(2\pi^1 J_{NC\alpha} T_{NC\alpha}) \bullet \cos(2\pi^2 J_{NC\alpha} T_{NC\alpha})$. Considering the average, upper, and lower range of $^1J_{NC\alpha}$ and $^2J_{NC\alpha}$ determined by Wirmer and Schwalbe (2002), we used values of $^2J_{NC\alpha}$ = 9.0 Hz (orange traces), 7.7 Hz (purple traces), and 6.4 Hz (blue traces)

in creating the plots. Because $^{1}J_{NC\alpha}$ and $^{2}J_{NC\alpha}$ coupling constants were also shown to be correlated, the green points show intersects at those values. (A) $T_{NC\alpha}$ (the constant time period divided by two) set to 23.5 ms. All ^{1}J values in the reported range (shaded region) result in observable signal, in excess of +5% for resonances with small ^{1}J values. (B) T set to 24.5 ms. Resonances with small ^{1}J couplings are completely suppressed, and any residual signals (less than 4%) from large values are inverted. (C) T set to 25.5 ms. All ^{1}J values in the reported range result in inverted signals, with up to 6% remaining intensity.

Further Pulse Sequence Details:

Unless indicated, pulses are applied along the *x*-axis. All narrow (wide) pulses are applied with a flip angle of 90° (180°). Hard pulses for ^{1}H, ^{15}N, and ^{13}C are centered at 4.7, 120, and 56 ppm, respectively. Selective $^{13}C'$ ($^{13}C^{\alpha}$) sinc shaped pulses are centered at 174 ppm (56 ppm) and a null 113 ppm away. Delays used are: Δ = 2.4 ms; δ = 0.11 ms. Phase cycling: ϕ_1 = x,-x, ϕ_2 = x,x,-x,-x, ϕ_{rec} = x,-x. A 3-9-19 watergate pulse scheme is used in the reverse INEPT transfer to suppress the strong water signal. Garp1 decoupling is used on ^{15}N during acquisition. To accomplish States quadrature detection for the ^{15}N indirect dimension, a second FID is collected for each increment by changing the ϕ_1 phase to y,-y. For States-TPPI acquisition, the ϕ_1 and ϕ_{rec} phases are also incremented by 180° every other ^{15}N increment. Gradients are Würst shaped z-axis gradients with a length of 1 ms. Gradient strengths (G/cm): G_1: 5, G_2: 7, G_3: 17.

5. Backbone NMR Resonance Assignment of the Catalytic Subunit of cAMP-Dependent Protein Kinase A in complex with AMP-PNP

Larry R. Masterson, Lei Shi, Marco Tonelli, Alessandro Mascioni, Michael M. Mueller, and Gianluigi Veglia

Submitted to Biomolecular NMR Assignments October 2008

The catalytic subunit of protein kinase A is intimately involved with a number of signal transduction pathways and has been used as a benchmark to study the structural biology and biochemistry for the entire kinase family of enzymes. Here, we report the backbone assignment of the intact 41 kDa catalytic subunit bound to AMP-PNP.

Introduction

Protein kinase mediated phosphorylation is a key process for cellular signal transduction. Approximately 30% of all proteins in the human genome undergo reversible phosphorylation (1). As a consequence, abnormal protein kinase activity has been recognized as one of the main causes for diseases such as cancer, diabetes, rheumatoid arthritis, and some types of dilated cardiomyopathy. In fact, drugs such as rapamycin and flavopiridol have been shown to target protein kinase activity (2).

Since Protein Kinase A (EC: 2.7.11.11) is the most highly conserved protein kinase family member, it has served as the prototype to study the kinase family (3). The catalytic subunit of PKA (PKAc) shares the domain organization of all of the kinase catalytic domains, with a *bean-like* arrangement of two lobes to form an active site cleft. The small lobe is primarily associated with binding and positioning the nucleotide ligand, while the large lobe binds a protein substrate and provides a docking surface for the regulatory subunits (3). Highly conserved motifs within PKAc include: the C-helix (located in the small lobe) and a series of dynamic loops responsible for catalysis and positioned at the fringe between the small and the large lobe (the peptide positioning loop, the activation loop, the Mg^{2+}-positioning loop, Glycine-rich loop, and DFG loop). Mutations within these motifs have been shown to cause large scale changes in kinase activity and/or proper folding of the enzyme (3).

While a wealth of structural data is available for PKAc from x-ray crystallography (3), little is known about the atomic level detail of its dynamics or interaction with membrane proteins. Due to their dynamic nature, intact protein kinase

domains are notoriously resistant to NMR analysis and a handful of NMR studies are now available. Our initial NMR investigation on PKAc demonstrated the existence of a positive allosteric cooperativity driven in part by a dynamic sampling of conformational states. In order to analyze the dynamic transitions of PKAc by NMR, we have carried out the backbone resonance assignment of the enzyme bound to AMP-PNP (5'-adenylyl-β,γ-imidodiphosphate) using ^{2}H, ^{13}C, and ^{15}N labeled protein and extensive selective labeling. To date, we reached ~80% of the total resonances, leaving only regions in helices D and E (both located at the stable core of the enzyme) unassigned.

Methods and Experiments

PKAc expression and purification was carried out based on the protocols established by Taylor and co-workers (4). For triple resonance experiments, the perdeuterated protein was purified from cells were grown in M9 minimal media supplemented with $^{15}NH_4Cl$ and D-glucose-$^{13}C_6$-$^{2}H_7$ in 80% $^{2}H_2O$. Extensive selective amino acid labeling was utilized for the identification of 15 residue types in $^{1}H/^{15}N$ TROSY-HSQC spectra: Ala, Val, Leu, Ile, Lys, Arg, Thr, Ser, Gly, His, Phe, Tyr, Trp, Met, and Asn. Introduction of specific 1-^{13}C and ^{15}N labeled amino acids for dual amino acid labeling allowed the assignment of the following unique $^{13}C'$-^{15}N linked dipeptides in the primary sequences: LL (3 occurrences), LI (3 occurrences), FG (2 occurrences), and GF (1 occurrence). Cells were grown at 30 ºC until reaching an OD of 1.2, at which the temperature was reduced to 24 ºC and protein overexpression was induced by the addition of 0.5 mM IPTG (isopropyl-β-D-thiogalactopyranoside). Typically, expression was allowed to proceed for 12 hours (deuterated media) or 5

hours (non-deuterated media). Cells were harvested by centrifugation at 7,000 *g* for 20 minutes at 4 ºC and flash frozen in N_2 (*l*) prior to purification. Cell pellets were resuspended in lysis buffer containing 100 mM MES (pH 6.5), 5 mM β-mercaptoethanol, and 0.15 mg/ml lysozyme; and lysed on ice by sonication for 10 minutes. Cell debris were cleared by centrifugation at 75,000 *g* and PKAc was isolated by binding the supernatant batch-wise to P11 phosphocellulose resin and eluting over a gradient of 0-250 mM K_2HPO_4 dissolved in lysis buffer. Fractions containing purified enzyme were pooled and dialyzed against 20 mM K_2HPO_4 (pH 6.5), 25 mM KCl, and 5 mM β-mercaptoethanol (buffer A). Since recombinant PKAc is found to have three phosphoisoforms (containing either 2, 3, or 4 phosphoserine or phosphothreonine residues), isoforms were separated by cation exchange chromatography on a Mono-S column using a gradient of 0-30% KCl dissolved in buffer A. All NMR experiments were performed on the isoform containing three phosphates.

Triple resonance NMR experiments were performed on a sample containing 0.28 mM PKAc, while all other samples contained 0.5-1.0 mM total enzyme. Samples were dissolved in 180 mM KCl, 20 mM KH_2PO_4 (pH 6.5), 1 mM NaN_3, 20 mM DTT, 15 mM $MgCl_2$, 5% ^2H_2O, and 15 mM of the non-hydrolyzable nucleotide analogue, AMP-PNP. All experiments were performed at 300 K on Varian Inova 800 or 600 MHz NMR spectrometers equipped with triple resonance cryogenic probes with z-axis pulse field gradients. Backbone assignments were based on TROSY versions of HSQC, HNCA, HN(CO)CA, HNCAB, HNCO, HN(CA)CO (5), NOESY-HSQC (τ_{mix} = 100 ms) as well as a non-TROSY version of HSQC-NOESY-HSQC (τ_{mix} = 100 ms).

Unique resonances resulting from dual selective amino acid labeling were observed by using either a 2D HNCO or a CCLS-HSQC sequence (6). All data were processed using the software NMRPipe (7) and visualized with SPARKY (8).

Extent of assignments and data deposition

A $^{1}H/^{15}N$-TROSY-HSQC of perdeuterated PKAc is shown in Figure 1. Backbone resonance assignments have been achieved for 80% of the non-proline amide groups, ~80% of the C^{α} and C' resonances and 47% of the C^{β} resonances. Despite a handful of missing assignments which are located in two long helices in the stable core of the enzyme (9), all of the highly conserved regions have been assigned and include the C-helix, peptide positioning, activation, Mg^{2+}-positioning, glycine-rich, and DFG–loops. These resonance assignments compare well with the previous partial assignment (~55%) of the apo-enzyme (10). It should be noted that higher signal-to-noise was observed when PKAc was bound to AMP-PNP relative to the apo form of the enzyme, although the glycine and two phenylalanine resonances in the ^{184}DFGF^{187} sequence were broadened beyond detection. These were assigned using the apo enzyme with a CCLS-HSQC experiment as has been noted elsewhere (6).

The chemical shift values for ^{1}H, ^{13}C, and ^{15}N resonances of PKAc have been deposited at the BioMagResBank (http://www.bmrb.wisc.edu) under accession number 15985.

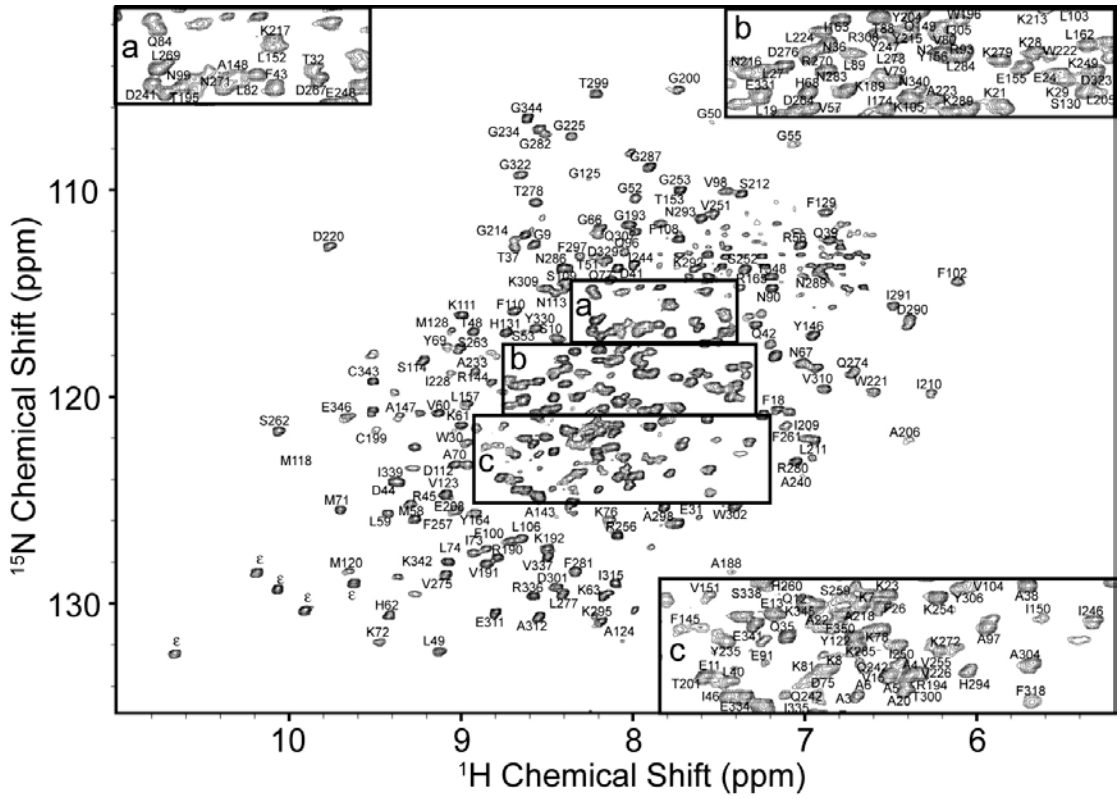


Figure 1: $^{1}H/^{15}N$-TROSY-HSQC spectrum of perdeuterated PKAc complexed to AMP-PNP in 20 mM KH_2PO_4 (pH 6.5), 180 mM KCl, 20 mM DTT, and 1 mM DTT collected at 300 K on a Varian Inova 800 MHz NMR spectrometer. Assignment information is placed next to peaks and subpanels a, b, and c are provided for clarity. Indolic nitrogen resonances of tryptophan are denoted with the symbol ε.

References

(1) Cohen P (2002) *Nat Cell Biol* 4, E127-30.

(2) Cohen P (2001) *Eur J Biochem* 268, 5001-5010.

(3) Johnson DA, Akamine P, Radzio-Andzelm E, Madhusudan M & Taylor SS (2001) *Chem Rev* 101, 2243-2270.

(4) Yonemoto WM, McGlone ML, Slice LW & Taylor SS (1991) *Methods Enzymol* 200, 581-596.

(5) Grzesiek S & Bax A (1992) *Journal of Magnetic Resonance* 96, 432-440.

(6) Tonelli M, Masterson LR, Hallenga K, Veglia G & Markley JL (2007) *J Biomol NMR* 39, 177-185.

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

(8) Goddard TD & Kneller DG (2006) SPARKY3.113, University of California, San Francisco.

(9) Yang J, Garrod SM, Deal MS, Anand GS, Woods J,Virgil L. and Taylor S (2005) *J Mol Biol* 346, 191-201.

(10) Langer T, Vogtherr M, Elshorst B, Betz M, Schieborr U, Saxena K and Schwalbe H (2004) *Chembiochem* 5, 1508-1516.

6. Allosteric Cooperativity in Protein Kinase A

Larry R. Masterson, Alessandro Mascioni, Nathaniel J. Traaseth, Susan S. Taylor, and Gianluigi Veglia

Reprinted with permission from the Proceedings of the National Academy of Sciences (2008) *105, 506-511.*

Allosteric signaling in proteins requires long-range communication by highly conserved residues, often triggered by ligand binding. In this paper, we map the allosteric network in the catalytic subunit of protein kinase A using nuclear magnetic resonance spectroscopy. We show that positive allosteric cooperativity is generated by nucleotide and subsrate binding during the transitions through the major conformational states: apo, intermediate, and closed. The allosteric network is disrupted by a single-site mutation (Y204A), which also decouples the cooperativity of ligand binding. Since protein kinase A is the prototype for the entire kinome, these findings may serve as a paradigm for describing long-range coupling in other protein kinases.

Introduction

The protein kinase superfamily is one of the largest gene families in eukaryotes, representing 2-4% of most genomes (1, 2). The human genome, for example, has over 500 predicted protein kinases. Protein kinases relay an extracellular signal into a biological response by phosphoryl transfer on specific substrates resulting in regulation of cell division, memory, differentiation, cell growth, and most other processes in the cell with exquisite precision. Since the first structure of the catalytic subunit of protein kinase A (PKA-C) was determined in 1991 (3, 4), there has been significant progress in filling the structural space of this family. Over 60 unique kinase structures are now available in the Protein Data Bank (www.rcsb.org). All of these structures share a highly conserved, kidney-shaped core that in PKA-C comprises residues 40-300. The core is comprised of a small, N-terminal lobe formed by β-strands which binds and positions ATP during catalysis, and a large lobe, which provides a docking surface for the substrate. The N-terminus ends with a single helix of approximately 40 residues (A helix). The 50 residues comprising the C-terminus wrap around the two lobes, docking into a hydrophobic pocket located in the small lobe (3, 4).

In spite of sharing a highly conserved core, protein kinases are remarkably diverse in terms of their regulation, activation, and substrate recognition. The crystal structures obtained suggest that kinases are highly dynamic with allosteric networks that radiate throughout the molecule. Unfortunately, crystal structures are static and often only include the kinase core. While there have been extensive studies of kinases in solution including hydrogen/deuterium exchange coupled with mass spectrometry

(H/D-MS), fluorescence anisotropy (5-7), and small angle X-ray scattering (8, 9), these results lack atomic level resolution and offer an incomplete picture about protein dynamics and recognition mechanisms. This presents a substantial limitation in understanding the enzymatic cycle of kinases, leaving many questions unanswered. How do these enzymes discriminate cognate and non-cognate substrates? What is the role of the nucleotide in substrate recognition? Does allostery play a fundamental role in these recognition processes?

This work investigated the transitions between the major states of the intact PKA-C (apo, intermediate, and closed) using NMR spectroscopy under functional conditions. This detailed residue-by-residue analysis revealed the existence of positive allosteric cooperativity which is triggered by ligand binding. In addition, NMR chemical shift mapping showed that a single mutation (Y204A) in the peptide positioning loop is able to disturb the allosteric network and decouple the cooperativity.

Results

NMR spectroscopy of the major conformationaz states.

The chemical shift changes in PKA-C upon ligand binding were monitored using 2D $^{1}H/^{15}N$ TROSY-HSQC (10) spectra. To mimick ATP, the non-hydrolyzable nucleotide AMP-PNP (aden-5'-yl imidodiphosphate) was used. The AMP-PNP-saturated enzyme was then titrated with Kemptide to mimic the Michaelis complex. **Figure 1** shows the spectra of the four distinct conformational states of the C-subunit:

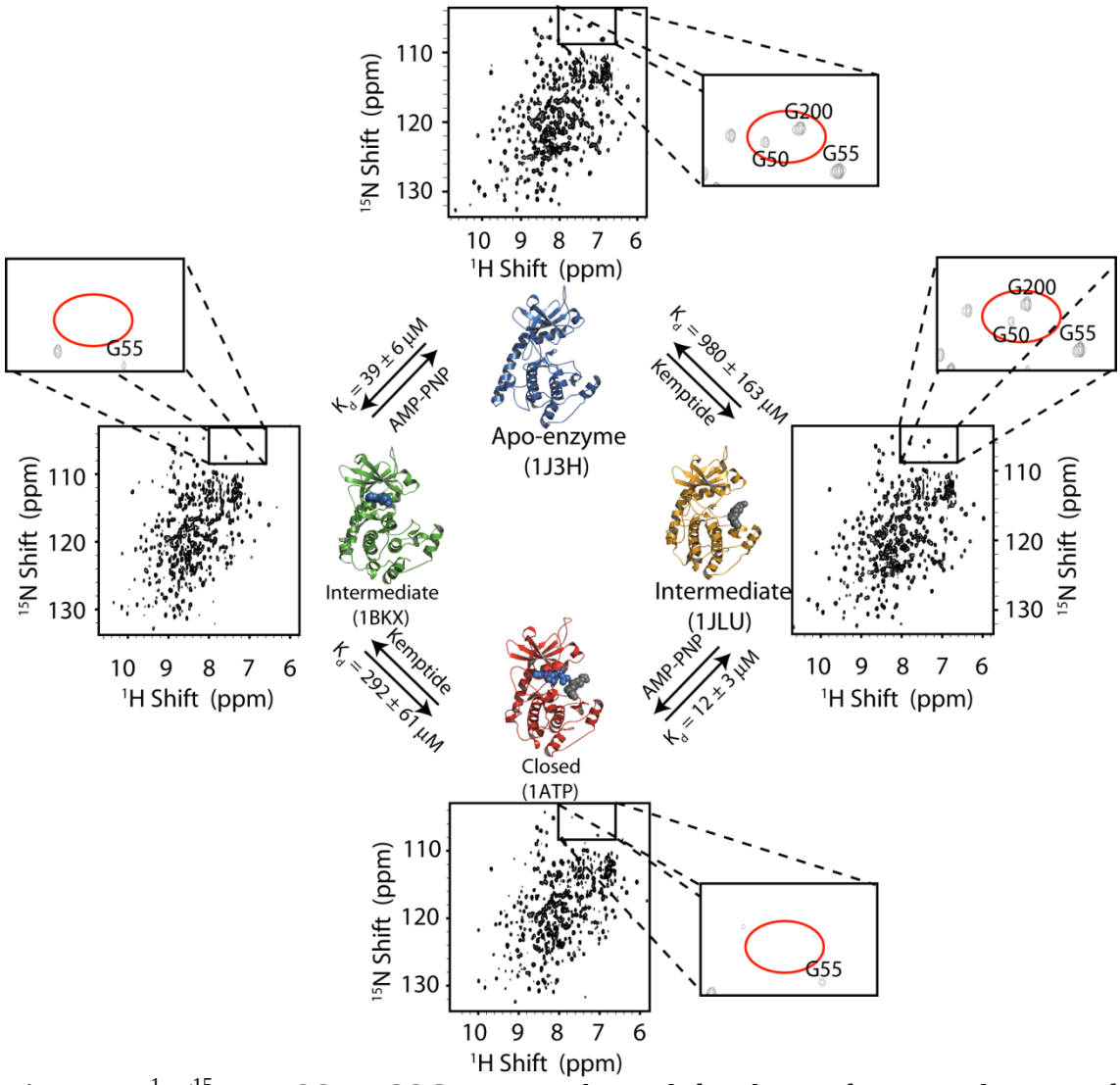


Figure 1: ^{1}H/^{15}N TROSY-HSQC spectra obtained for the conformational states of PKA-C: apo (top), intermediate-N (left), intermediate-S (right), and closed (bottom). Expanded boxes highlight residues of the glycine-rich and peptide positioning loops. Titrations using AMP-PNP first were performed with 0.63 mM C-subunit and titrations using Kemptide first were performed with 0.44 mM C-subunit.

apo, intermediate-N (enzyme saturated with nucleotide), intermediate-S (enzyme saturated with substrate), and closed. A total of 253 of 337 expected amide peaks were suitable for chemical shift perturbation ($\Delta\delta$) and lineshape analyses. These peaks are well distributed across the protein, giving an overall view of the enzyme's behavior during ligand binding. While the majority of the peaks displayed fast exchange upon ligand binding, several of the resonances were exchange broadened under saturating ligand conditions, supporting the existence of multiple conformational states and a complex energy landscape for this enzyme (11, 12).

In the following synopsis, we describe the changes in the fingerprints of the wild-type and Y204A mutant of PKA-C upon the addition of nucleotide and substrate.

AMP-PNP binding.

Figure 2 summarizes the chemical shift perturbations occurring at each amide site upon AMP-PNP binding. As expected for AMP-PNP binding, a cluster of residues within the small lobe of the enzyme is highly sensitive to the addition of nucleotide. The most pronounced changes ($\Delta\delta > 0.05$ ppm) are observed in residues of the β3-strand and the highly conserved glycine-rich loop (β1- and β2-strands), which actively participate in positioning the nucleotide. Crystal structures have shown that the β3-strand flanks the nucleotide binding site and contains K72 which coordinates the α– and β–phosphates of ATP (13, 14). Gradual decreases of $\Delta\delta$ towards the C- and N-terminal regions from the glycine-rich loop suggest that nucleotide binding affects distal sites

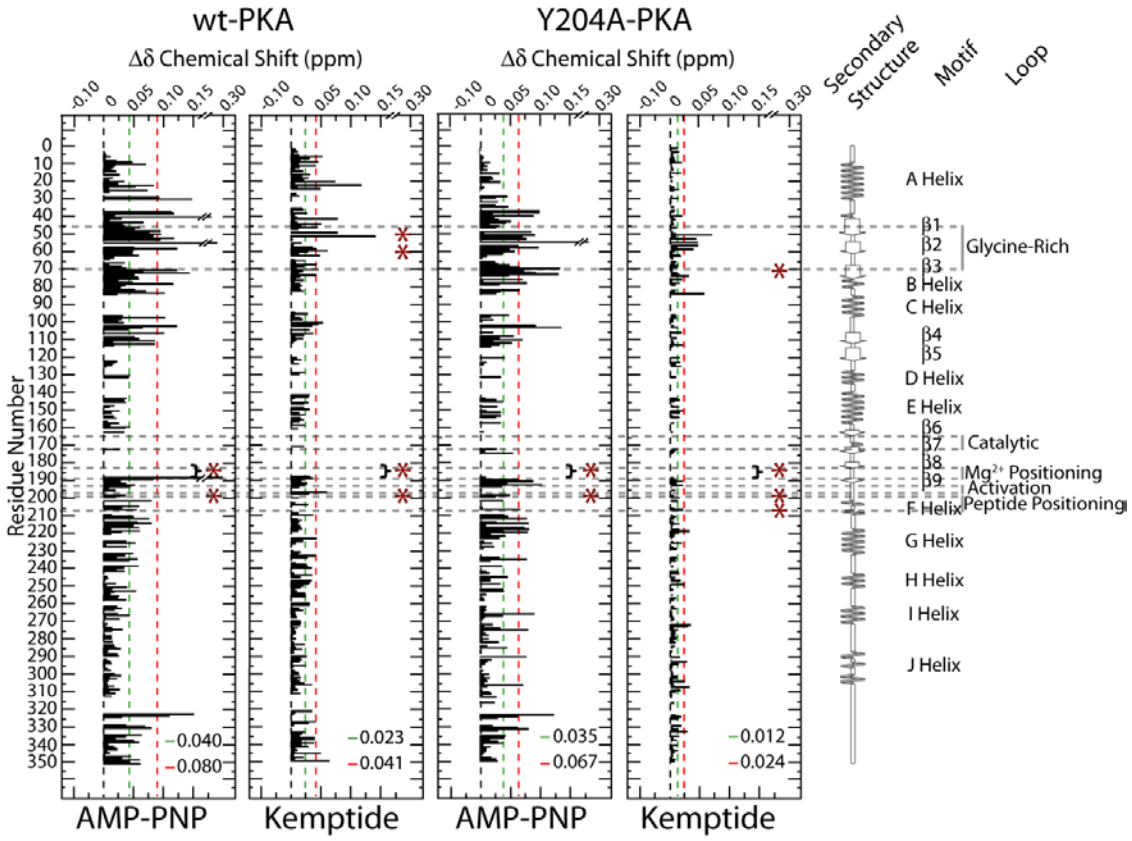


Figure 2: Residue specific perturbations observed during ligand binding. Histograms show the combined $^{1}H/^{15}N$ chemical shift perturbations ($\Delta\delta = \sqrt{\Delta\delta_H{}^2 + (0.154\Delta\delta_N)^2}$) versus residue. The cutoffs for <Δδ> and <Δδ> plus one standard deviation are given as green and red dashed lines, respectively. Asterisks indicate exchange broadened residues.

within the small lobe of the enzyme. This is evident by the large Δδ values for residues that are >35 Å away from helices A and B in the crystal structure (**Figure 3**).

A second region is centered at A188, which is located at the surface of the large lobe where substrate docking would take place **(Figures 2** and **3**). This large perturbation is important, since crystal structures reveal that the A188 amide group points toward the substrate recognition sequence (15). This region contains highly conserved loops, including the DFG-loop which has been proposed to be a switch between active and inactive kinase conformations (16). The resonances belonging to the DFG-loop (F185, G186, and F187) were broadened beyond detection (see asterisks in **Figure 2**), indicating conformational interconversion between multiple states, which were silent to the x-ray analysis. Significant changes are observed for residues T201 (~0.24 ppm) and G200 (which became exchange broadened), which are conserved throughout the Ser/Thr kinase family (17). The large perturbations at the peptide positioning loop, with concomitant changes in the DFG- and glycine-rich loops, reveal that the nucleotide affects both the small and large lobes, perhaps priming the enzyme for substrate binding. This is supported by x-ray studies carried out by Taylor and co-workers that show ligand induced conformational changes at these domains (18). The F helix, which is located near the peptide positioning loop, also has large Δδ values which gradually decrease toward the C-terminus where a third region is sensitive to binding. The large Δδ values for residues at the C-terminus reflect the structural features of the enzyme highlighted in **Figure 3**: the C-terminus wraps around the small lobe of the

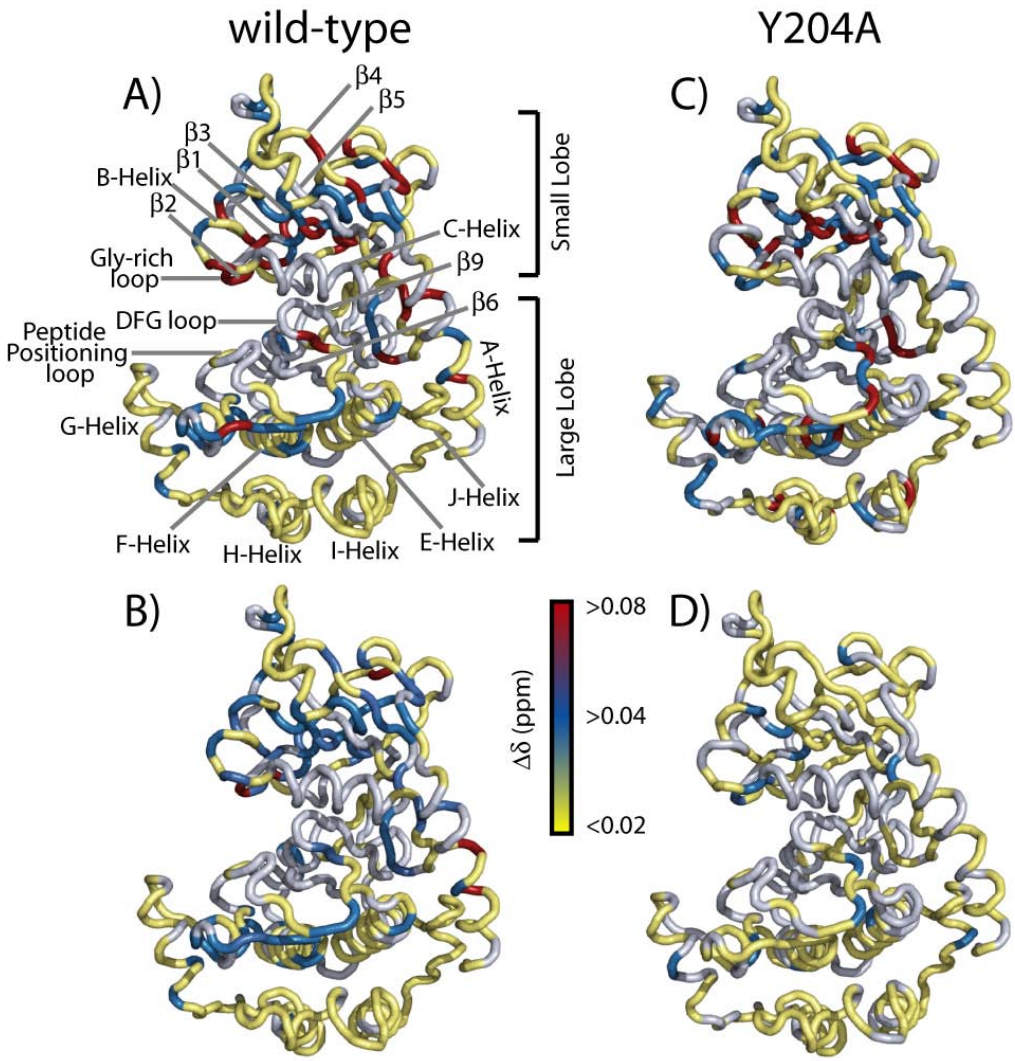


Figure 3: Chemical shift perturbation mapping. Perturbations were mapped for the nucleotide bound (intermediate-N) and closed forms of wild-type (left) and Y204A (at right) enzymes.

enzyme and positions F325-E337 near the glycine-rich loop (19). Marked changes along the primary sequence in the large lobe (including the F, G, H, I, and J helices) suggest that nucleotide binding radiates its effects to distal sites of both lobes.

Kemptide binding.

After saturation with the nucleotide (i.e. formation of intermediate-N), Kemptide was titrated into the sample to mimic the closed state (i.e. Michaelis complex). While AMP-PNP caused large $\Delta\delta$ throughout the enzyme (~30% of the assigned residues above have $\Delta\delta > 0.05$), the perturbations observed upon binding Kemptide are significantly lower ($<\Delta\delta> = 0.023$), with less than 8% of the residues perturbed more than 0.05 ppm. Residue specific differences are evident in **Figure 2**, but the overall pattern of $\Delta\delta$ for the Kemptide titration is similar to that of the nucleotide. The most prominent perturbations are localized in three regions: centered in the glycine-rich loop, between the catalytic and the peptide positioning group, and in the C-terminal portion of the enzyme. The largest perturbation occurs at S53, which in crystal structures bridges the nucleotide and substrate. The backbone amide interacts with the γ-phosphate of ATP, and its side chain hydroxyl group forms a hydrogen bond with the backbone carbonyl of the P-site residue of the substrate (15). A gradual decrease of $\Delta\delta$ was observed from the glycine-rich loop toward the N- and C-termini. Significant $\Delta\delta$ at distal sites of the small lobe again suggest sensitivity to binding, but here they are induced by changes at the surface of the large lobe.

Interestingly, Kemptide addition does not quench the exchange broadening caused by nucleotide binding for residues in the DFG- and peptide positioning loops. Additionally, G50 and G52 became exchange broadened, pointing to the existence of interconverting conformations. As with AMP-PNP, the extent of $\Delta\delta$ decreases from the F to J helices. Small yet significant $\Delta\delta$ values may be indicative of a redistribution of the conformational energy upon substrate binding, an effect that was not anticipated from x-ray structures. Of course, Kemptide is a small substrate and the effects described could be further widespread or increased with larger substrates.

Reverse Order of Ligand Binding.

The titrations were repeated by inverting the order of ligand additions, adding Kemptide first (i.e. intermediate-S), and then saturating the enzyme with AMP-PNP to form the Michaelis complex (the histograms with $\Delta\delta$ versus residue number are provided in the Supporting Information **Figure S1**). Unlike AMP-PNP addition to the apo-enzyme, saturation with Kemptide caused relatively modest changes to the enzyme fingerprint. Only the phosphorylated S10 located in the A helix shows a drastic chemical shift change. However, smaller changes are detected throughout the entire enzyme's backbone, which indicates that Kemptide binding radiates its effects to both lobes. Subsequent addition of AMP-PNP to the Kemptide-bound enzyme causes changes in three regions that are similar to the direct titration: around the glycine-rich loop, between the catalytic and peptide positioning loops, and the unstructured C-terminal portion of the enzyme (**Figure S1**). Irrespective of the order of ligands used in the titrations, the 2D $^{1}H/^{15}N$ TROSY-HSQC spectra of the closed form of the enzyme

are superimposable, which illustrates the thermodynamic equivalence for binding pathways.

Cooperativity of ligand binding.

Resonances undergoing fast exchange were selected for the calculation of the dissociation constants (K_d). **Table 1** summarizes the calculated K_d values and representative fits are provided in **Figure S2** of the Supporting Information. We measured a K_d of 39 ± 6 μM for AMP-PNP and an apparent K_d value of 292 ± 61 μM for Kemptide, which compare well with values reported for ATP (K_d = 25 ± 1 μM) and Ala-Kemptide (LRRAALG) (K_d = 230 ± 70 μM) attained by transient state kinetic studies (20). The relative affinity of substrate and nucleotide was also examined from the reverse order titration using Kemptide first. In this case, the K_d for Kemptide was 980 ± 163 μM and the apparent K_d for AMP-PNP was 12 ± 3 μM. Thus, binding of the first ligand appears to enhance the affinity of the second ligand, which corresponds to positive cooperativity.

Disruption of the allosteric network and decoupling positive cooperativity.

Is it possible to remove the positive cooperativity by disrupting the allosteric network? To answer this question, we repeated the NMR titrations with the Y204A-PKA-C mutant. This mutation is located in the peptide positioning loop, and causes the enzyme to bind ATP and a peptide inhibitor in a similar manner as the wild-type enzyme (21). However, solvent accessibility and thermostability studies have pointed to a disruption

of the allosteric network, which likely causes the 30-fold decrease in the enzyme's turnover rate (22, 23). The direct titration using the nucleotide first resulted in a K_d value of 78 ± 20 μM, and an apparent K_d value of 834 ± 92 μM for Kemptide. The

Table 1: Binding constants (*Kd*) measured from the NMR titrations of C-subunit PKA.

	AMP-PNP First		Kemptide First	
	$K_d^{\text{AMP-PNP}}$	K_d^{Kemptide}	K_d^{Kemptide}	$K_d^{\text{AMP-PNP}}$
wt-PKA	39 ± 6	292 ±61	980 ± 163	12 ± 3
Y204A-PKA	78 ± 20	834 ± 92	700 ± 105	68 ± 4

reverse titration (Kemptide first) resulted in a K_d value of 700 ± 105 μM for Kemptide and an apparent K_d value of 68 ± 4 μM for AMP-PNP. The K_d values (**Table 1**) for the mutant show no cooperative binding effect, indicating that the two binding events have become decoupled.

Figure 2 shows the Δδ upon the direct titration of AMP-PNP, while **Figure S1** shows the corresponding reverse titration. The overall trend of Δδ for AMP-PNP is similar to the wild-type, with three regions primarily affected: the glycine-rich loop, peptide positioning loop, and C-terminus. However, these perturbations are attenuated and altered in comparison to the wild-type enzyme. In particular, smaller changes were observed at the distal sites near the N-and C- terminal regions of the small lobe, residues 100-110, and the residues surrounding the peptide positioning loop of the large lobe. **Figure 3** illustrates two different allosteric networks of interactions between the wild-type and mutant enzymes. While the allosteric network surrounds the active site of the wild-type enzyme, the allosteric effects induced by the nucleotide binding in Y204A are not localized and dispersed throughout the large lobe (see Δδ for helices G, H, I, J). The disruption is more apparent for the binding of Kemptide, where modest chemical shift changes throughout the protein backbone are observed. The only region showing significant perturbation is the glycine-rich loop, which is drastically attenuated in comparison to the wild-type enzyme. Moreover, the effects that radiated to distal sites in both the small and large lobes are no longer present for Kemptide (**Figure 3**).

Lineshape Analysis for wild-type PKA-C.

Among all of the resolved peaks of the enzyme, we isolated those resonances experiencing an intermediate two-site exchange during these titrations and carried out a quantitative lineshape analysis to evaluate the binding kinetics. Since K_d values were independently measured from residues affected by fast exchange, we calculated both the *on*- and *off*- rates ($K_d=k_{off}/k_{on}$) by analyzing the NMR line broadening dependence from k_{on}[ligand] and k_{off} (24). The lineshape analyses of 5 different peaks (F110, F108, F100, R190, and R93), resulted in *on*-rates for AMP-PNP ranging between 0.7 and 2.8 x 10^6 M^{-1} s^{-1}. Examples of lineshape fitting and a table reporting the k_{on} values for 5 residues are reported in the Supporting Information (**Figure S3** and **Table S1**). These values are very similar to those reported for PKA-C using transient kinetic measurements (20). As for other enzymes (24, 25), k_{on} is approximately two orders of magnitude lower than the bimolecular-diffusion controlled rates (~10^8-10^9 M^{-1} s^{-1}). The low *on*-rate supports the hypothesis of a rather complex binding mechanism for the nucleotide, which requires large conformational changes in the small lobe (18). In fact, it has been hypothesized that segments within the small lobe are not only responsible for binding, aligning, and positioning the nucleotide for the formation of a catalytically competent complex, but changes in both lobes occur to properly position the substrate, and shield the active site from bulk water (26).

Discussion

NMR spectroscopy is becoming the method of choice to analyze allostery (27-35), cooperativity (27, 32, 36-38), energy landscapes (11, 32, 37-39), and coordinate of reactions for small and large enzymes (11, 38, 40). From the analysis of our NMR titrations emerges the first vivid picture of the complex changes occurring during the transitions from apo to intermediate, and from intermediate to closed conformations in the backbone of PKA-C. Specifically, the NMR characterization of allostery resulted in the identification of positive cooperativity for ligand binding. A key result in our data is the presence of exchange broadening for highly conserved residues in the kinase family (i.e. G50, G52, F185, G186, and G200) under saturating conditions of nucleotide or nucleotide plus substrate (see **Figures 1, 2,** and **S2**). This suggests that PKA-C interconverts between multiple dynamic or conformational states. Thus, it is possible that in solution the conformational states identified by x-ray crystallography are in equilibrium. In the absence of ligands, the protein favors the conformation observed in the crystal structure of the apo state, with small populations of the other states. The addition of ligand alters this equilibrium to favor the intermediate and closed conformations. In this manner, the first ligand can drive the system close to the final state, facilitating binding of the second ligand.

To test the existence of the allosteric network, we mutated Y204 to Ala. Y204 is located in the P+1 loop and is part of the tetrad of residues that keeps the F helix, D helix, and P-2 binding site together through electrostatic interactions (21, 23). This single mutation caused disruptions in the allosteric network and removed the binding cooperativity. The comparison of our NMR titrations between the wild-type and Y204A

mutant enzymes shows that this allostery is mediated by both local (contiguous) and long-range (non-contiguous or disperse) (41) changes. During the binding of nucleotide, we observed a cluster of hydrophobic residues (L59, A70, M120, and L106) that were particularly sensitive ($\Delta\delta$ ~0.1 ppm) in the wild-type enzyme. These residues make contact with the evolutionarily conserved 'spine' (L106, L95, F185, Y164) recently identified by Ten Eyck and co-workers (42), to form a hydrophobic lining about the adenine ring (**Figure 4**). Sensing and positioning of the nucleotide at this region could act as a trigger for signaling which is propagated across a network of residues away from the active site. Perturbations greater than 0.05 ppm are traceable across conserved residues within van der Waals radii from the hydrophobic lining (**Figure 4**). Several contiguous pathways of intramolecular cross-talk can also be traced from the hydrophobic lining to residues which can propagate signals to and from the glycine-rich, activation, DFG, and peptide positioning loops at the entrance of the active site. We also detected non-contiguous propagation for a small subset of residues in distal sites, for which no clear pathway of interaction can be traced. In the wild-type, two regions display relatively large $\Delta\delta$ values. The first is located at the N-terminus (~24 Å from the nucleotide binding pocket). The second region, located at ~12 Å from the nucleotide binding pocket, comprises F100 and F102, which bridge the small and large lobes (18) and demonstrates that the signal is transmitted across the lobes of the enzyme. A possible explanation for non-contiguous signaling is that small local conformational changes add up and culminate in crucial regions of the enzyme, providing a long-range intramolecular signaling mechanism up to ~24 Å from the nucleotide binding site (see red residues in **Figure 4**). In contrast, the Y204A mutation

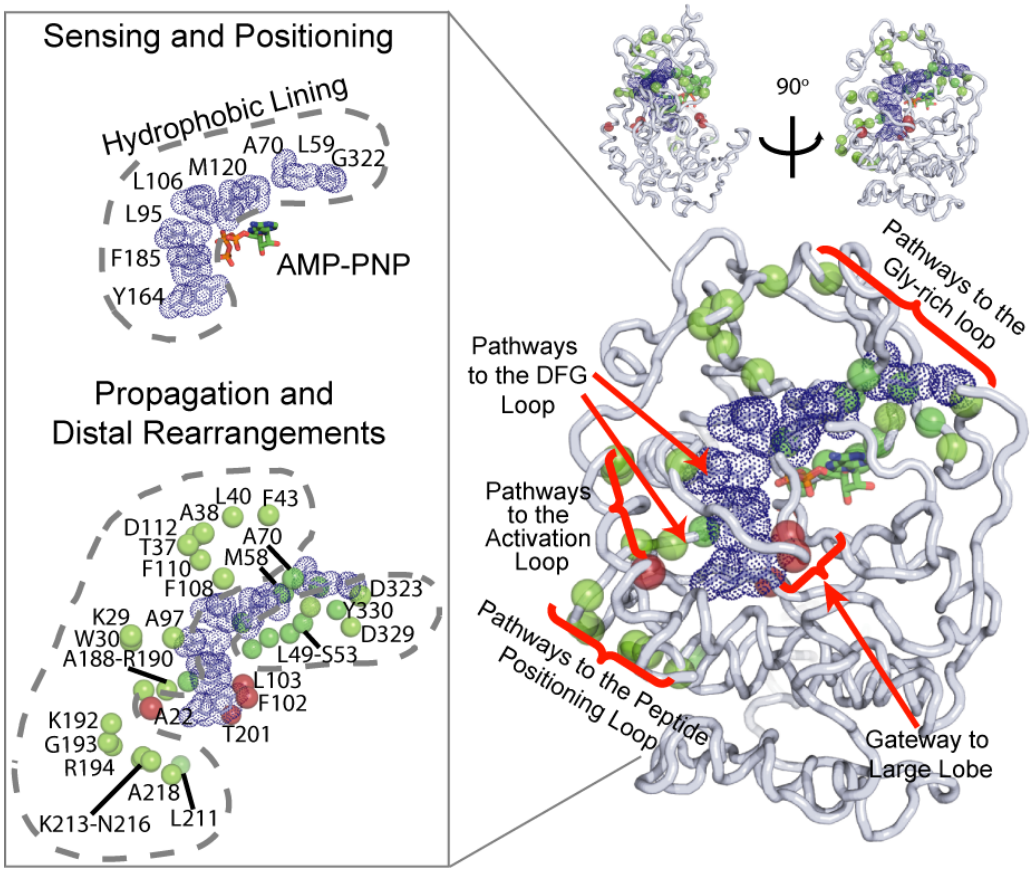


Figure 4: Proposed model for the allosteric network of interactions in wild-type PKA-C. Blue: residues of the hydrophobic lining. Green: residues showing the contiguous network of interactions located within van der Waals radii from the hydrophobic lining. Red: residues showing non-contiguous interactions located outside of the van der Waals contacts.

causes an overall rearrangement of the allosteric communication between the two lobes of the enzyme upon nucleotide binding and an interruption of this communication upon substrate binding.

But to what extent do the chemical shift changes reflect the conformational transitions to the states observed by x-ray crystallography? Chemical shift changes encompass several different phenomena (conformational interconversion, different binding modes, binding kinetics, electrostatics, and hydrogen bonding) (43-46), and in the absence of spin relaxation measurements to supplement these results, it is difficult to quantify the different contributions. However, we found a direct correlation between the measured $\Delta\delta$ values and the difference in the Cα atomic coordinates between the apo (1J3H) and intermediate (1BKX) structures and between the intermediate (1BKX) and closed (1ATP) structures (**Figure S4**). The overall trend is remarkably similar, which supports the hypothesis that allostery is *partially* mediated by conformational changes throughout the entire enzyme. These observations are in agreement with other studies carried out using fluorescence anisotropy (18). However, this does not exclude the role of protein dynamics for the positive allosteric cooperativity. In fact, the crystal structures for the closed forms of Y204A and wild-type PKA-C are identical (21), although solution data indicate Y204A is more dynamic (23). This increased dynamics may contribute to the disruption of the allosteric network and decoupling of the binding cooperativity. Further analyses of NMR spin relaxation properties supplemented by solvent accessibility and thermostability are needed to determine the relative importance of dynamics and conformational changes in the positive allosteric cooperativity of the C-subunit.

The modern view of molecular allostery recognizes the existence of a dynamic structural ensemble of the protein native states (47-51). Our data suggest the reshaping of the allosteric energy landscape through ligand binding: the population of states which are encoded in the apo enzyme shift in the presence of ligand to generate positive cooperativity. This work demonstrates that PKA-C is an exquisite example of allostery and communication networks between distal regions. These finding can serve as a paradigm for describing long-range coupling in other protein kinases.

Methods

Sample preparation.

PKA-C expression and activity assays were performed as previously described (52, 53) (details are provided in the Supplementary Information and **Figure S5**). NMR samples consisted of 0.44 and 0.63 mM of wild-type C-subunit for the forward and reverse titrations, respectively. Samples for the mutant enzyme were ~0.24 mM for the forward and reverse titrations. All samples were prepared in 20 mM KH_2PO_4, 10 mM DTT, 180 mM KCl, 10 mM $MgCl_2$, 1 mM NaN_3, and 5% ^2H_2O.

NMR resonance assignments and ligand titrations.

NMR experiments were carried out using Varian Inova spectrometers operating at either 800.24 or 599.71 MHz and equipped with triple resonance cryoprobes. Backbone resonance assignments for the wild-type enzyme were carried out using HNCA, HN(CO)CA, HNCACB, HNCO, and HN(CA)CO experiments on $^2H/^{15}N/^{13}C$ labeled PKA samples at 300 K. Assignments were also assisted by selective ^{15}N amino acid

labeling. These samples were titrated with the ligands in order to determine assignments for the bound states. **Tables S2** and **S3** summarize all the experiments and the acquisition parameters. Our resonance assignments compare well with the previous partial (55%) assignment obtained by Schwalbe and co-workers (54). Spectra for the Y204A samples were virtually superimposable with resonances of wild-type enzyme, thus assignments in this case were based on comparison (overlay of spectra are provided in **Figure S6** of the supplementary information). NMR data were processed and visualized using the software NMRPipe (55) and SPARKY (56). Chemical shift perturbations from NMR titrations were calculated with the formula (**Eqn. 1**):

$$\Delta\delta = \sqrt{\Delta\delta_H{}^2 + (0.154\,\Delta\delta_N)^2} \qquad (1)$$

where $\Delta\delta$ is the combined chemical shift, and $\Delta\delta_H$ and $\Delta\delta_N$ are the differences of ^1H and ^{15}N chemical shifts, respectively, between the first and last point of the titrations. The weighting factor for the ^{15}N chemical shift was calculated as previously described (24).

Assuming a 1:1 enzyme:ligand ratio (3, 4), K_d values were determined by a non-linear fit (**Eqn. 2**) of resonances in fast exchange between free enzyme and bound enzyme (57):

$$\Delta\Omega_i = \frac{K_d + [L_T] + [E_T] - \sqrt{(K_D + [L_T] + [E_T])^2 - 4[L_T][E_T]}}{2[E_T]}\Delta\Omega \qquad (2)$$

where $[E]_T$ and $[L]_T$ are total concentrations of enzyme and ligand added, respectively, and $\Delta\Omega_i$ and $\Delta\Omega$ are the chemical shift differences between the observed and initial resonances and the final and initial resonances, respectively. Non-linear least-squares fitting of the quadratic equation to the data was performed using Mathematica software

(58). The calculation for the dissociation constants in the Michaelis complex which accounts for the small amount of free enzyme are described in the Supporting Information.

References

(1) Manning G, Whyte DB, Martinez R, Hunter T & Sudarsanam S. (2002) *Science* 298: 1912-1934.

(2) Tchieu JH, Fana F, Fink JL, Harper J, Nair TM, Niedner RH, Smith DW, Steube K, Tam TM, Veretnik S, *et al*. (2003) *Nucleic Acids Res* 31: 342-344.

(3) Knighton DR, Zheng J, Eyck LFT, Xuong NH, Taylor SS & Sowadski JM. (1991) *Science* 253: 414-420.

(4) Knighton DR, Zheng JH, Ten Eyck LF, Ashford VA, Xuong NH, Taylor SS & Sowadski JM. (1991) *Science* 253: 407-414.

(5) Iyer GH, Garrod S, Woods VL,Jr & Taylor SS. (2005) *J Mol Biol* 351: 1110-1122.

(6) Zhang J, Ma Y, Taylor SS & Tsien RY. (2001) *Proc Natl Acad Sci U S A* 98: 14997-15002.

(7) Burns-Hamuro LL, Ma Y, Kammerer S, Reineke U, Self C, Cook C, Olson GL, Cantor CR, Braun A & Taylor SS. (2003) *PNAS* 100: 4072-7.

(8) Vigil D, Blumenthal DK, Taylor SS & Trewhella J. (2006) *J Mol Biol* 357: 880-889.

(9) Taylor SS, Kim C, Vigil D, Haste NM, Yang J, Wu J & Anand GS. (2005) *Biochim Biophys Acta* 1754: 25-37.

(10) Pervushin K, Riek R, Wider G & Wuthrich K. (1997) *Proc Natl Acad Sci U S A* 94: 12366-71.

(11) Boehr DD, McElheny D, Dyson HJ & Wright PE. (2006) *Science* 313: 1638-1642.

(12) Boehr DD, Dyson HJ & Wright PE. (2006) *Chem Rev* 106: 3055-3079.

(13) Taylor SS, Radzio-Andzelm E, Madhusudan , Cheng X, Eyck LFT & Narayana N. (1999) *Pharmacol. Ther.* 82: 133-141.

(14) Iyer GH, Moore MJ & Taylor SS. (2005) *J Biol Chem* 280: 8800-8807.

(15) Aimes RT, Hemmer W & Taylor SS. (2000) *Biochemistry* 39: 8325-8332.

(16) Taylor SS, Haste NM & Ghosh G. (2005) *Cell* 122: 823-825.

(17) Burns-Hamuro LL, Barraclough DM & Taylor SS. (2004) *Methods Enzymol* 390: 354-374.

(18) Johnson DA, Akamine P, Radzio-Andzelm E, Madhusudan M & Taylor SS. (2001) *Chem Rev* 101: 2243-2270.

(19) Narayana N, Cox S, Shaltiel S, Taylor SS & Xuong N. (1997) *Biochemistry* 36: 4438-4448.

(20) Lew J, Taylor SS & Adams JA. (1997) *Biochemistry* 36: 6717-6724.

(21) Yang J, Ten Eyck LF, Xuong NH & Taylor SS. (2004) *J Mol Biol* 336: 473-487.

(22) Moore MJ, Adams JA & Taylor SS. (2003) *J Biol Chem* 278: 10613-10618.

(23) Yang J, Garrod SM, Deal MS, Anand GS, Woods J,Virgil L. & Taylor S. (2005) *J Mol Biol* 346: 191-201.

(24) Tugarinov V & Kay LE. (2003) *J Mol Biol* 327: 1121-33.

(25) Fersht A. (1999) *Structure and Mechanism in Protein Science,* (W.H. Freeman & Co., New York),

(26) Gerstein M, Lesk AM & Chothia C. (1994) *Biochemistry* 33: 6739-6749.

(27) Linse S & Chazin WJ. (1995) *Protein Sci* 4: 1038-1044.

(28) Maler L, Blankenship J, Rance M & Chazin WJ. (2000) *Nat Struct Biol* 7: 245-250.

(29) Volkman BF, Lipson D, Wemmer DE & Kern D. (2001) *Science* 291: 2429-33.

(30) Kern D, Zuiderweg, E.R. (2003) *Curr Opin Struct Biol* 13: 748-57.

(31) Eisenmesser EZ, Millet O, Labeikovsky W, Korzhnev DM, Wolf-Watz M, Bosco DA, Skalicky JJ, Kay LE & Kern D. (2005) *Nature* 438: 117-121.

(32) Leung DW & Rosen MK. (2005) *Proc Natl Acad Sci U S A* 102: 5685-5690.

(33) Popovych N, Sun S, Ebright RH & Kalodimos CG. (2006) *Nat Struct Mol Biol* 13: 831-838.

(34) Ferguson AD, Amezcua CA, Halabi NM, Chelliah Y, Rosen MK, Ranganathan R & Deisenhofer J. (2007) *Proc Natl Acad Sci U S A* 104: 513-518.

(35) Das R, Esposito V, Abu-Abed M, Anand GS, Taylor SS & Melacini G. (2007) *Proc Natl Acad Sci U S A* 104: 93-98.

(36) Stevens SY, Sanker S, Kent C & Zuiderweg ER. (2001) *Nat Struct Biol* 8: 947-952.

(37) Das R, Abu-Abed M & Melacini G. (2006) *J Am Chem Soc* 128: 8406-8407.

(38) Velyvis A, Yang YR, Schachman HK & Kay LE. (2007) *Proc Natl Acad Sci U S A* 104: 8815-8820.

(39) Stevens SYS,S., Kent C & Zuiderweg ER. (2001) *Nat Struct Biol.* 8: 947-52.

(40) Sprangers R, Gribun A, Hwang PM, Houry WA & Kay LE. (2005) *Proc Natl Acad Sci U S A* 102: 16678-16683.

(41) Clarkson MW, Gilmore SA, Edgell MH & Lee AL. (2006) *Biochemistry* 45: 7693-7699.

(42) Kornev AP, Haste NM, Taylor SS & Ten Eyck LF. (2006) *Proc Natl Acad Sci USA* 103: 17783-17788.

(43) Sitkoff D & Case DA. (1997) *J Am Chem Soc* 119: 12262-12273.

(44) Revington M, Holder TM & Zuiderweg ER. (2004) *J Biol Chem* 279: 33958-33967.

(45) Matsuo H, Walters KJ, Teruya K, Tanaka T, Gassner GT, Lippard SJ, Kyogoku Y & Wagner G. (1999) *J Am Chem Soc* 121: 9903-9904.

(46) Hellstern S, Pegoraro S, Karim CB, Lustig A, Thomas DD, Moroder L & Engel J. (2001) *J Biol Chem* 276: 30845-52.

(47) Freire E. (1999) *Proc Natl Acad Sci U S A* 96: 10118-10122.

(48) Kumar S, Ma B, Tsai CJ, Sinha N & Nussinov R. (2000) *Protein Sci* 9: 10-19.

(49) Hammes GG. (2002) *Biochemistry* 41: 8221-8228.

(50) Hammes-Schiffer S & Benkovic SJ. (2006) *Annu Rev Biochem* 75: 519-541.

(51) Swain JF & Gierasch LM. (2006) *Curr Opin Struct Biol* 16: 102-108.

(52) Yonemoto WM, McGlone ML, Slice LW & Taylor SS. (1991) *Methods Enzymol* 200: 581-596.

(53) Cook PF, Neville ME, Vrana KE, Hartl FT & Roskoshi R. (1982) *Biochemistry* 21: 5794-5799.

(54) Langer T, Vogtherr M, Elshorst B, Betz M, Schieborr U, Saxena K & Schwalbe H.(2004) *Chembiochem* 5: 1508-1516.

(55) Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J & Bax A. (1995) *J Biomol*

NMR 6: 277-293.

(56) Goddard TD & Kneller DG. (2006) 3.113, *SPARKY*.

(57) Johnson PE, Tomme P, Joshi MD & McIntosh LP. (1996) *Biochemistry* 35: 13895- 13906.

(58) Wolfram Research I. (2005) *Mathematica.*

Supplementary Information

Sample Preparations

^{15}N-labeled enzyme used the titrations was expressed in 2 liters of ^{15}N-enriched minimal media using $^{15}NH_4Cl$ as the sole source of nitrogen. Samples used for resonance assignments were grown in 80% $^{2}H_2O$ media containing ^{2}H,^{13}C-labeled glucose and $^{15}NH_4Cl$ as the sole sources of nitrogen and carbon. All of the NMR sample preparations were tested for activity using gel shift assays, coupled enzyme assays, and direct ^{31}P detection of the phosphorylation reaction in the NMR tube (see **Figure S5**). The C-subunit expressed from *E. coli* contained three different phospho-isoforms (Slice LW & Taylor SS. 1989, *J Biol Chem* **264**: 20940). The most abundant isoform, phosphorylated at T197, S338, and S10, was used in this study. The k_{cat} and K_M values obtained for the wild type enzyme using Kemptide (LRRASLG) as a substrate compare well with the current literature (data not shown) (Lew J, Taylor SS & Adams JA., 1997, *Biochemistry* **36**: 6717-6724).

Selective amino acid labeling of the wild type C-subunit with ^{15}N Ala, Arg, Phe, Gly, Ile, Lys, Leu, Met, Asn, Ser, Thr, Val, and Trp were used to both aid in the assignments of triple resonance spectra and to discern assignments in each state by titrating these samples with ligand. To unambiguously assign Leu-Leu, Leu-Ile, Phe-Gly, and Gly-Phe pairs present in the primary sequence, samples were selectively labeled with ^{15}N (amide) and ^{13}C (carbonyl) amino acids (**Table S2**). These

connectivities were assigned using a $^{1}H/^{15}N$-2D version of the TROSY-HNCO experiment or using a CCLS-HSQC (Tonelli, M, et al, 2007 *Journal of Biomolecular NMR*, **39**: 177-185).

Table S1: Values obtained from NMR line-shape analysis of five residues undergoing intermediate chemical exchange. Lines-shapes were simulated using the model for two-site exchange to determine the off-rate of AMP-PNP binding. Using the calculated K_D value from residues undergoing F/FI exchange, the on-rates were then obtained as reported below.

Residue	**k_{on} ($M^{-1}s^{-1}$)**
F110 (8.67 ppm, 115.7 ppm)	1.1×10^6
F108 (7.81 ppm, 111.6 ppm)	0.70×10^6
F100 (8.86 ppm, 127 ppm)	0.83×10^6
R190 (8.85 ppm, 128.9 ppm)	2.8×10^6

Table S2: NMR acquisition parameters used for the resonance assignment and titration experiments of the C-subunit of PKA-C. Unless noted otherwise, all experiments were performed at 800 MHz ^{1}H field strength. All multidimensional data were acquired with 2048 complex points in the ^{1}H dimension, 80 complex points in the ^{15}N dimension, and 80 complex points in the ^{13}C dimension. Spectra were processed with a sine-squared bell function shifted by 90^{o} and zero filled to double the number of points in each dimension prior to Fourier transformation.

	Number of Scans	**Number of Increments* (^{13}C spectral width)**	**Number of Increments* (^{15}N spectral width)**	**Experiment Time**
HNCA	40	40 (7600 Hz)	40 (2500 Hz)	156 hrs
HN(CO)CA	40	40 (7600 Hz)	40 (2500 Hz)	156 hrs
HNCACB	40	40 (11322 Hz)	40 (2500 Hz)	157 hrs
HNCO	16	40 (3770 Hz)	40 (2500 Hz)	63 hrs
HN(CA)CO	64	32 (8291 Hz, 600 MHz)	40 (2100, 600 MHz)	192 hrs
^{1}H/^{15}N-HSQC	40	-	40 (2500 Hz)	2 hrs

Table S3: Double selective amino acid labeling schemes and corresponding assignments obtained for the C-subunit of PKA. The labeling scheme used in these experiments incorporates ^{15}N at the amide position and ^{13}C at the carbonyl position.

Labeling Pattern	**^{13}C,^{15}N-Leu/^{15}N-Ile**	**^{13}C-Phe/^{15}N-Gly**	**^{13}C-Gly/^{15}N-Phe**
Assignment Obtained	I163, I176, I228, L173, L269, L273	G55 and G185	F187

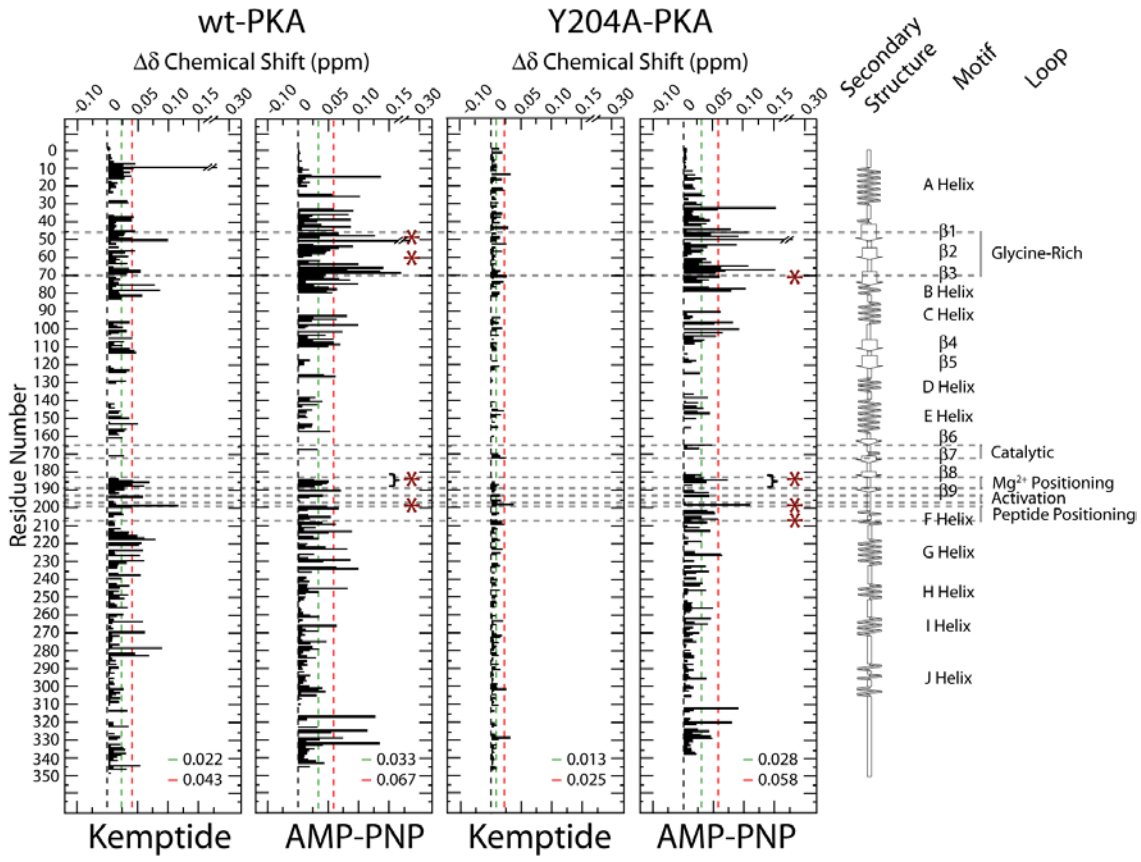


Figure S1: Histograms summarizing the 1H/15N combined chemical shift perturbations (Δδ) for the reverse titration (Kemptide titration to saturation first, followed by AMP-PNP). Cutoffs for the <Δδ> and <Δδ> plus one standard deviation are given as green and red dashed lines, respectively. Asterisks indicate exchange broadened residues detected at the end point of the titration.

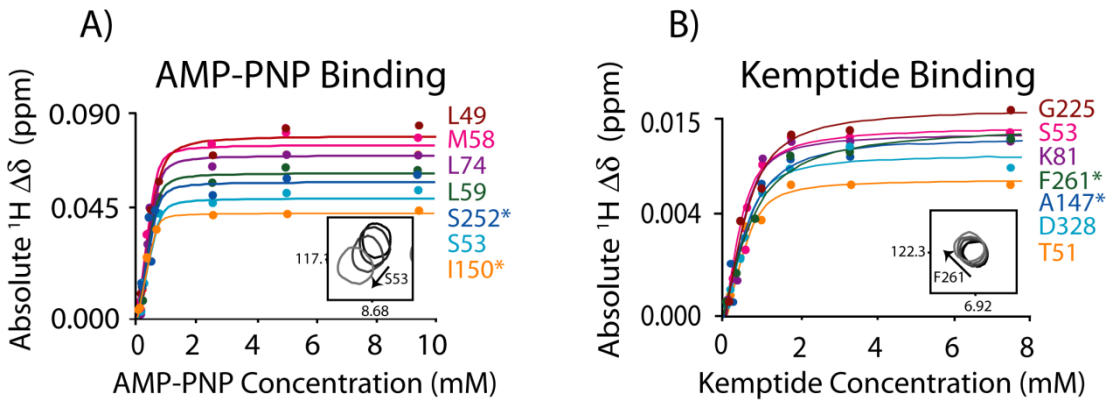


Figure S2: Binding curve fits and representative spectra for titrations using AMP-PNP and Kemptide. Long-range allosteric effects are shown by asterisks in which the residues shown are greater than 12 Å (AMP-PNP binding) or 20 Å (substrate binding) as measured in the crystal structure 1ATP. Measurements of the K_d values reported in this study only included residues within an 8 Å sphere from the ligand.

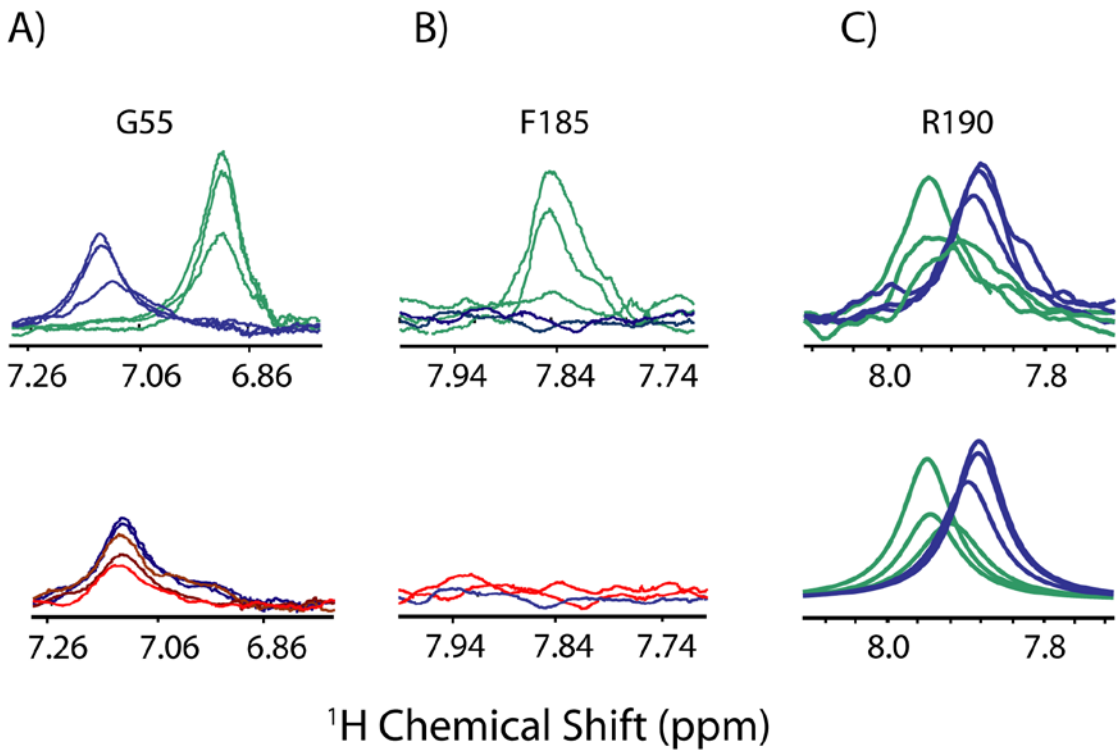


Figure S3: Representative one-dimensional 1H NMR chemical shift changes occurring in the A) glycine-rich loop and B) the DFG-loop. Green, blue, and red traces represent resonances corresponding to the *apo*, *intermediate*, and *closed* states, respectively. The changes during the Kemptide titration in panel A is followed by a gradient going from blue to red. During the titration of AMP-PNP into PKA, some resonances exchanged on the intermediate timescale between *apo* and AMP-PNP bound states (panels **A** and **C**). Under this timescale of chemical exchange, it is possible to simulate the NMR line-shape for the determination of k_{on}. An example of the line-shape fit, performed with R190, is shown in the bottom panel of **C**.

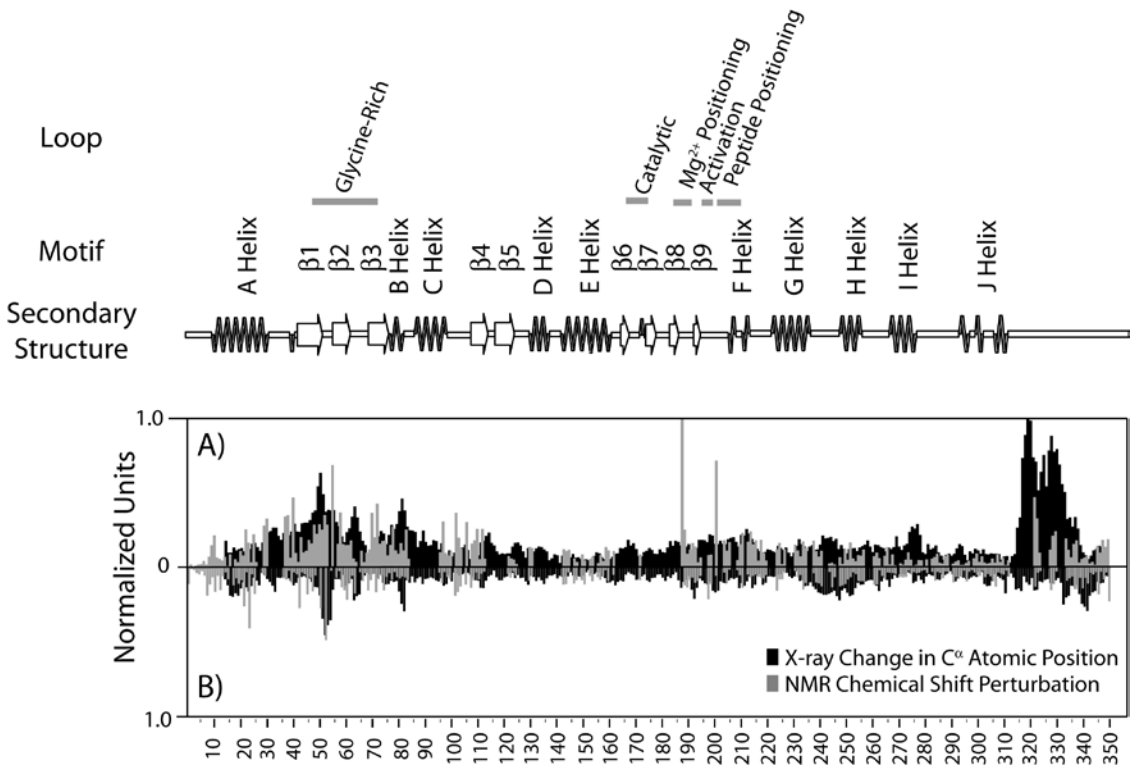


Figure S4: Comparison of NMR chemical shifts perturbations and differences in x-ray structures of C-subunit PKA in different conformational states. The changes predicted by x-ray (black bars) were calculated as the difference of C^{α} coordinates between two different states using the following PDB files: 1J3H (apo), 1BKX (intermediate), and 1ATP (closed). The PDB coordinates were superimposed using the most invariant parts of the structures as determined by hydrogen/deuterium exchange (49). **A)** Comparison of $\Delta\delta$ obtained upon titration of the apo form of the C-subunit of PKA with AMP-PNP (gray) with the difference between the Cα coordinates of the intermediate and apo states; and **B)** comparison of $\Delta\delta$ obtained upon the addition of Kemptide (gray) to AMP-PNP bound enzyme with the difference between the Cα coordinates of closed and intermediate states.

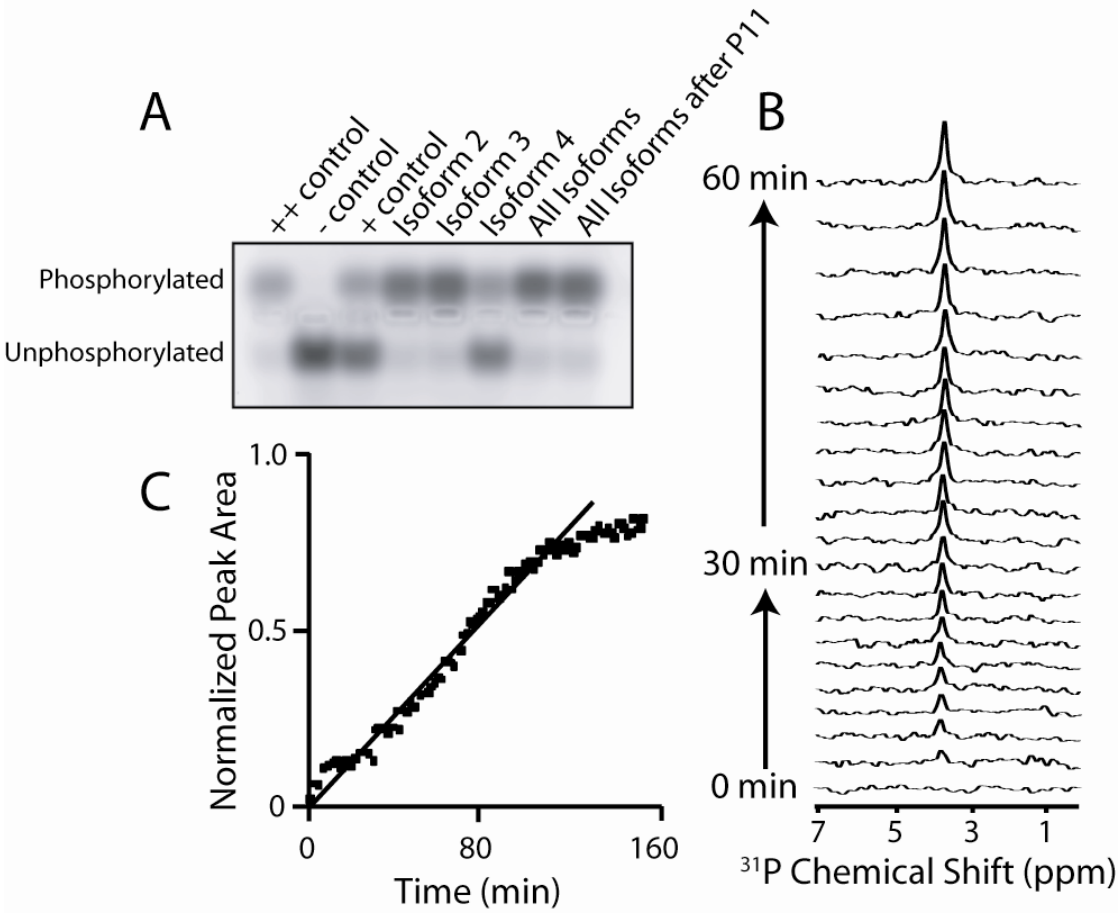


Figure S5: Activity of the C-subunit during purification and under NMR conditions. **A)** Agarose gel shifts for phosphorylated and unphosphorylated products after incubation with C-subunit PKA in the presence of MgATP. The activity of C-subunit at different stages of the purification is shown. Commercial PKA was used as a control at different concentrations (inidicated by + and ++) and recombinant phospho-isoforms containing 2, 3, and 4 phosphates are indicated in each lane. **B)** Time-dependent ^{31}P NMR spectra of for the appearance of the phospho-Kemptide product during catalytic phosphorylation by 2-15 μM C-subunit PKA. The measurements under these conditions

included 100 mM ATP and 1-10 mM Kemptide in 50 mM MOPS at pH 6.5. **C)** Plot of the normalized peak area of the phosphoserine resonance in Kemptide at each time increment. The linear portion up to ~80 minutes is expected for a pseudo-first order reaction under saturating concentrations (>100-fold excess) of ATP and Kemptide.

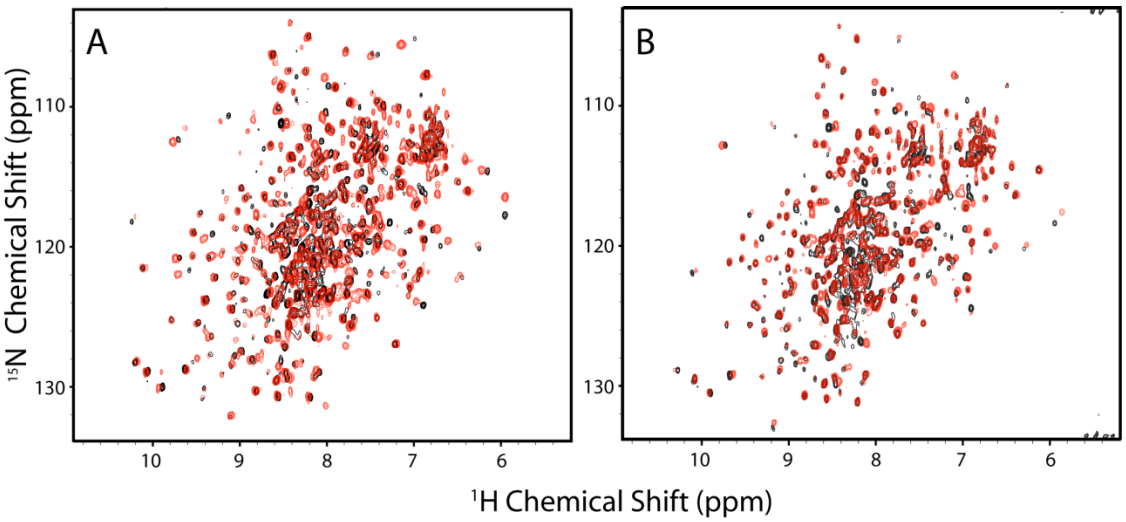


Figure S6: ^{1}H/^{15}N TROSY-HSQC overlay for the wild type C-subunit of PKA (red) and the mutant Y204A (black) in the **A)** the apo and **B)** the AMP-PNP bound states. Since very little perturbation occurred to the resonance positions of Y204A upon the addition Kemptide, the spectra were are similar to the above for the direct Kemptide titration (spectra A) or the titration of Kemptide to the AMP-PNP bound state (B).

Measurement of the apparent dissociation constants for the Michaelis complex (Kemptide and AMP-PNP)

The branched pathway mechanism for enzyme (E), nucleotide (L), and peptide (P) is given below. This is needed to account for the small free enzyme concentration at the end of the first titration before the second addition of either ligand or peptide.

$$\begin{aligned} EL &\Leftrightarrow E + L \; K_{d1} & K_{d1} &= \frac{[E][L]}{[EL]} \\ EP &\Leftrightarrow E + P \; K_{d2} & K_{d2} &= \frac{[E][P]}{[EP]} \\ ELP &\Leftrightarrow EP + L \; K_{d1'} & K_{d1'} &= \frac{[EP][L]}{[ELP]} \\ ELP &\Leftrightarrow EL + P \; K_{d2'} & K_{d2'} &= \frac{[EL][P]}{[ELP]} \end{aligned} \qquad (1)$$

The branched pathway equilibrium shown above cannot be solved explicitly for both dissociation constants, $K_{d1'}$ and $K_{d2'}$. To calculate $K_{d1'}$, we need to solve the following equation,

$$K_{d1'} = \frac{[EP][L]}{[ELP]} \qquad (2)$$

The free ligand concentration (Eqn. 3) and the [ELP] concentration (Eqn. 4) are substituted into Eqn. 2 to give Eqn. 5.

$$[L] = [L_T] - [EL] - [ELP] - [EP] \qquad (3)$$

$$[ELP] = [E_T] f_B \qquad (4)$$

$$K_{d1'} = \frac{[EP]([L_T]-[EL]-[E_T]f_B-[EP])}{[E_T]f_B} \quad (5)$$

Next we solve equations involving K_{d1} and K_{d2} in Eqn. 1 for [EP] and [EL]. These equations are shown below in an expanded form.

$$K_{d1} = \frac{[E][L]}{[EL]} = \frac{([E_T]-[EL]-[EP]-[ELP])([L_T]-[EL]-[EP]-[ELP])}{[EL]} \quad (6)$$

$$K_{d2} = \frac{[E][P]}{[EP]} = \frac{([E_T]-[EL]-[EP]-[ELP])([P_T]-[EL]-[EP]-[ELP])}{[EP]} \quad (7)$$

To simplify the solution to these equations, two assumptions are made: 1) $[P_T] >> [E_T]$ and 2) $[E_T] >> [EL]$. These assumptions are made since the enzyme was saturated with peptide before adding the nucleotide. Eqns. 6 and 7 simplify to,

$$K_{d1} = \frac{([E_T]-[EP]-[ELP])([L_T]-[EP]-[ELP])}{[EL]} \quad (8)$$

$$K_{d2} = \frac{([E_T]-[EP]-[ELP])[P_T]}{[EP]} \quad (9)$$

Eqns. 8 and 9 are then easily solvable for [EP] and [EL]. After inserting these solutions into Eqn. 5, the final expression for $K_{d1'}$ is:

$$K_{d1'} = \frac{P_T(f_B-1)K_{d1}([E_T]f_B-[L_T])}{f_B(P_TK_{d1}+K_{d2}(K_{d1}+[E_T]-[E_T]f_B))} \quad (10)$$

The fraction bound (f_B) is related to the observed chemical shifts,

$$f_B = \frac{\Delta\Omega_i}{\Delta\Omega} \quad (11)$$

where $\Delta\Omega_i$ and $\Delta\Omega$ are the chemical shift differences between the observed and initial position of resonances and the final and initial position of resonances, respectively. Substituting Eqn. 11 into Eqn. 10, and solving for $\Delta\Omega_i$, gives the final expression, which depends on $[L_T]$, $[P_T]$, $[E_T]$, K_{d1}, K_{d2}, $K_{d1'}$ and $\Delta\Omega$.

$$\Delta\Omega_i = \frac{\Delta\Omega}{2([P_T]K_{d1} + K_{d2}K_{d1'})[E_T]}([P_T]K_{d1}K_{d1'} + K_{d1}K_{d2}K_{d1'} + [P_T][E_T]K_{d1} + [E_T]K_{d2}K_{d1'} - \sqrt{-4[P_T][L_T][E_T]K_{d1}([P_T]K_{d1} + K_{d2}K_{d1'}) + K_{d2}K_{d1'}(K_{d1} + [E_T]) + [P_T]K_{d1}(K_{d1'} + [L_T] + [E_T])^2})$$

(12)

A similar treatment is used to calculate $K_{d2'}$.

Mathematica software was used to fit Eqn. 12 to obtain $K_{d1'}$. All dissociation constant values are given in Table S4.

Bibliography

Abdulaev NG, *et al* (2006) *Biochemistry (N Y)* 45, 12986-12997.

Abergel D & Palmer AG (2004) *Chemphyschem* 5, 787-793.

Adams JA (2001) *Chem Rev* 101, 2271-2290.

Aimes RT, Hemmer W & Taylor SS. (2000) *Biochemistry* 39: 8325-8332.

Bax A & Ikura M (1991) *Journal of Biomolecular NMR* 1, 99-104.

Betz, M.; Saxena, K.; Schwalbe, H. (2006) *Curr. Opin. Chem. Biol. 10*, 219-225.

Boehr DD, McElheny D, Dyson HJ & Wright PE. (2006) *Science* 313: 1638- 1642.

Boehr DD, Dyson HJ & Wright PE. (2006) *Chem Rev* 106: 3055-3079.

Breitenlechner CB, *et al* (2004) *Biochemistry* 43, 7743-7749.

Brittsan AG & Kranias EG (2000) *J Mol Cell Cardiol* 32, 2131-9.

Buck, B., Zamoon, J., Kirby, T.L., DeSilva, T.M., Karim, C., Thomas, D.D., Veglia, G. (2003) *Protein Expr. Purif.* 30, 253-261.

Burns-Hamuro LL, Barraclough DM & Taylor SS. (2004) *Methods Enzymol* 390: 354-374.

Burns-Hamuro LL, Ma Y, Kammerer S, Reineke U, Self C, Cook C, Olson GL, Cantor CR, Braun A & Taylor SS. (2003) *PNAS* 100: 4072-7.

Carr SA, Biemann K, Shoji S, Parmelee DC & Titani K (1982) *Proc Natl Acad Sci U S A* 79, 6128-6131.

Clarkson, J.; Campbell, I. D. (2003) *Biochem. Soc. Trans. 31*, 1006-1009.

Clarkson MW, Gilmore SA, Edgell MH & Lee AL. (2006) *Biochemistry* 45: 7693-7699.

Clore GM & Gronenborn AM (1994) *Meth Enzymol* 239, 349-363.

Cohen P (2002) *Nat Cell Biol* 4, E127-30.

Cohen P (2001) *Eur J Biochem* 268, 5001-5010.

Cook PF, Neville ME, Vrana KE, Hartl FT & Roskoshi R. (1982) *Biochemistry* 21: 5794-5799.

Dalvit C (1992) *Journal of Magnetic Resonance (1969)* 97, 645-650.

Das R, Abu-Abed M & Melacini G. (2006) *J Am Chem Soc* 128: 8406-8407.

Das R, Esposito V, Abu-Abed M, Anand GS, Taylor SS & Melacini G. (2007) *Proc Natl Acad Sci U S A* 104: 93-98.

De Furia, F.G., Miller, D.R., Cerami, A., Manning, J.M. (1972) *J. Clin. Invest.* 51, 566-574.

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

DeWitt MM, MacLeod HM, Soliven B & McNally EM (2006) *J Am Coll Cardiol* 48, 1396-1398.

Downing AK (2004) *Protein NMR techniques*. Human Press, New Jersey.

Dolja,V.V., Peremyslov, V.V., Keller, K.E., Martin, R.R., Hong, J. (1998) *Virology* 252, 269-274.

Eisenmesser EZ, Millet O, Labeikovsky W, Korzhnev DM, Wolf-Watz M, Bosco DA, Skalicky JJ, Kay LE & Kern D. (2005) *Nature* 438: 117-121.

Evenas J,Evenäs J, Tugarinov V, Skrynnikov NR, . Goto NK, Muhandiram R, Kay LE (2001) *J Mol Biol* 309, 961-974.

Ferguson AD, Amezcua CA, Halabi NM, Chelliah Y, Rosen MK, Ranganathan R & Deisenhofer J. (2007) *Proc Natl Acad Sci U S A* 104: 513-518.

Fersht A. (1999) *Structure and Mechanism in Protein Science*, (W.H. Freeman & Co., New York),

Fiaux, J.; Bertelsen, E. B.; Horwich, A. L.; Wuthrich, K. (1995) *J. Biomol. NMR 5*, 339-44.

Fiaux J, Bertelsen EB, Horwich AL & Wuthrich K (2002) *Nature* 418, 207-211.

Freire E. (1999) *Proc Natl Acad Sci U S A* 96: 10118-10122.

Gardner KH & Kay LE (1998) *Annu Rev Biophys Biomol Struct* 27, 357-406.

Gerstein M, Lesk AM & Chothia C. (1994) *Biochemistry* 33: 6739-6749.

Graves, J.D., Krebs, E.G. (1999) *Pharmacol. Ther.* 82, 111-121.

Grzesiek S & Bax A (1992) *Journal of Magnetic Resonance* 96, 432-440.

Goddard, T.D., Kneller, D.G. (2006) SPARKY version 3.113, University of California, San Francisco.

Golovanov, A. P.; Blankley, R. T.; Avis, J. M.; Bermel, W. (2007) *J. Am. Chem. Soc.129*, 6528-6535.

Haghighi K, Haghighi K, Kolokathis F, Gramolini AO, Waggoner JR, Pater L, Lynch RA, Fan GC, Tsiapras D, Parekh RR, Dorn GW 2nd, MacLennan DH, Kremastinos DT, Kranias EG.*et al* (2006) *Proc Natl Acad Sci U S A* 103, 1388-1393.

Hammes GG. (2002) *Biochemistry* 41: 8221-8228.

Hammes-Schiffer S & Benkovic SJ. (2006) *Annu Rev Biochem* 75: 519-541.

Hansen DF & Led JJ (2003) *J Magn Reson* 163, 215-227.

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

Hauer, J.A., Barthe, P., Taylor, S.S., Parello, J., Padilla, A. (1999) *Protein Sci.* 8, 545-553.

Hellstern S, Pegoraro S, Karim CB, Lustig A, Thomas DD, Moroder L & Engel J. (2001) *J Biol Chem* 276: 30845-52.

Hu, J., Qin, H., Li, C., Sharma, M., Cross, T.A., Gao, F.P. (2007) *Protein Sci.* 16, 2153-2165.

Ikura M, Kay LE & Bax A (1990) *Biochemistry* 29, 4659-4667.

Iyer GH, Garrod S, Woods VL,Jr & Taylor SS. (2005) *J Mol Biol* 351: 1110-1122.

Johnson PE, Tomme P, Joshi MD & McIntosh LP (1996) *Biochemistry* 35, 13895-13906.

Johnson, D.A., Akamine, P., Radzio-Andzelm, E., Madhusudan, M., Taylor, S.S. (2001) *Chem. Rev.* 101, 2243-2270.

Kainosho M & Tsuji T (1982) *Biochemistry* 21, 6273-6279.

Kay LE, Ikura M, Tschudin R & Bax A (1990) *Journal of Magnetic Resonance* 89, 496-514.

Kay, LE (2005) *J. Magn. Reson. 173*, 193-207.

Kern D, Zuiderweg, E.R. (2003) *Curr Opin Struct Biol* 13: 748-57.

Knighton DR, Zheng JH, Ten Eyck LF, Ashford VA, Xuong NH, Taylor SS & Sowadski JM. (1991) *Science* 253: 407-414.

Knighton DR, Zheng J, Eyck LFT, Xuong NH, Taylor SS & Sowadski JM. (1991) *Science* 253: 414-420.

Kornev AP, Haste NM, Taylor SS & Ten Eyck LF. (2006) *Proc Natl Acad Sci USA* 103: 17783-17788.

Kumar S, Ma B, Tsai CJ, Sinha N & Nussinov R. (2000) *Protein Sci* 9: 10-19.

Langer T, Vogtherr M, Elshorst B, Betz M, Schieborr U, Saxena K and Schwalbe H (2004) *Chembiochem* 5, 1508-1516.

Langer T, Sreeramulu S, Vogtherr M, Elshorst B, Betz M, Schieborr U, Saxena K, Schwalbe H.*et al* (2005) *FEBS Lett* 579, 4049-4054.

Leung DW & Rosen MK. (2005) *Proc Natl Acad Sci U S A* 102: 5685-5690.

Lew J, Taylor SS & Adams JA. (1997) *Biochemistry* 36: 6717-6724.

Lian LY & Middleton DA (2001) *Progress in Nuclear Magnetic Resonance Spectroscopy* 39, 171-267.

Linse S & Chazin WJ. (1995) *Protein Sci* 4: 1038-1044.

Lundstrom, P.; Teilum, K.; Carstensen, T.; Bezsonova, I.; Wiesner, S.; Hansen, D. F.; Religa, T. L.; Akke, M.; Kay, L. E. (2007) *J. Biomol. NMR 38*, 199-212.

Mac Lennan DH & Kranias EG (2003) *Nature Reviews* 4, 566-577.

Maler L, Blankenship J, Rance M & Chazin WJ. (2000) *Nat Struct Biol* 7: 245-250.

Manning G, Whyte DB, Martinez R, Hunter T & Sudarsanam S. (2002) *Science* 298: 1912-1934.

Masterson, L.R., Mascioni, A., Traaseth, N.J., Taylor, S.S., Veglia, G. (2008) *Proc. Natl. Acad. Sci.* 105, 506-511.

Material Safety Data Sheet Triton X-100 (2005) MSDS, MSDS No. SLT2452, Science Lab, Houston, Texas.

Matsuo, H.; Walters, K. J.; Teruya, K.; Tanaka, T.; Gassner, G. T.; Lippard, S. J.; Kyogoku, Y.; Wagner, G. (1999) *J. Am. Chem. Soc 121*, 9903-9904.

McIntosh LP & Dahlquist FW (1990) *Q Rev Biophys* 23, 1-38.

Moore MJ, Adams JA & Taylor SS. (2003) *J Biol Chem* 278: 10613-10618.

Mori S, Abeygunawardana C, Johnson MO & van Zijl PC (1995) *J Magn Reson B* 108, 94-98.

Narayana N, Cox S, Shaltiel S, Taylor SS & Xuong N. (1997) *Biochemistry* 36: 4438-4448.

Simmerman HK, Kobayashi YM, Autry JM & Jones LR (1996) *Journal Biological Chemistry* 271, 5941-5946.

Sitkoff D & Case DA. (1997) *J Am Chem Soc* 119: 12262-12273.

Schmitt JP, *et al* (2003) *Science* 299, 1410-3.

Slice LW & Taylor SS (1989) *J Biol Chem* 264, 20940-20946.

Taylor SS, Yang J, Wu J, Haste NM, Radzio-Andzelm E, Anand G. (2004) *Biochim Biophys Acta* 1697, 259-269.

Taylor, S.S., Knighton, D.R., Zheng, J., Ten Eyck, L.F., Sowadski, J.M. (1992) *Annu. Rev. Cell. Biol.* 8, 429-462.

Walsh DA, Perkins JP & Krebs EG (1968) *J Biol Chem* 243, 3763-3765.

Wang Y & Jardetzky O (2002) *Protein Sci* 11, 852-861.

Weigelt J (1998) *J Am Chem Soc* 120, 10778-10779.

Wuthrich K (1986) *NMR of Proteins and Nucleic Acids.* John Wiley and Sons, New York.

Opella, S.J., Ma, C., Marassi, F.M. (2001) *Methods Enzymol.* 339, 285-313.

Parker MJ, Aulton-Jones M, Hounslow AM & Craven CJ (2004) *J Am Chem Soc* 126, 5020-5021.

Pervushin K, Riek R, Wider G & Wuthrich K (1997) *Proc Natl Acad Sci U S A* 94, 12366-71.

Pitto Venters, R. A.; Huang, C. C.; Farmer, B. T.,2nd; Trolard, R.; Spicer, D.; Fierke, C. A. (1995) *J Biomol NMR 5*, 339-44.

Popovych N, Sun S, Ebright RH & Kalodimos CG. (2006) *Nat Struct Mol Biol* 13: 831-838.

Revington M, Holder TM & Zuiderweg ER (2004) *J Biol Chem* 279, 33958-33967.

Riek R, Fiaux J, Bertelsen EB, Horwich AL & Wuthrich K (2002) *J Am Chem Soc* 124, 12144-12153.

Riek, R.; Wider, G.; Pervushin, K.; Wuthrich, K. (1999) *Proc. Natl. Acad. Sci. USA 96*, 4918-4923.

Schmitt,J.P., Kamisago, M., Asahi, M., Li, G.H., Ahmad, F., Mende, U., Kranias, E.G., MacLennan, D.H., Seidman, J.G., Seidman, C.E. (2003)*Science* 299, 1410-3.

Simmerman, H.K., Jones, L.R. (1998) *Physiol. Rev.* 78, 921-47.

Sitkoff D & Case DA (1997) *J Am Chem Soc* 119, 12262-12273.

Stevens SY, Sanker S, Kent C & Zuiderweg ER. (2001) *Nat Struct Biol* 8: 947-952.

Sprangers R, Gribun A, Hwang PM, Houry WA & Kay LE. (2005) *Proc Natl Acad Sci U S A* 102: 16678-16683.

Sprangers, R.; Kay, L. E. (2007) *Nature 445*, 618-622.

Swain JF & Gierasch LM. (2006) *Curr Opin Struct Biol* 16: 102-108.

Taylor SS, Radzio-Andzelm E, Madhusudan , Cheng X, Eyck LFT & Narayana N. (1999) *Pharmacol. Ther.* 82: 133-141.

Taylor SS, *et al* (2004) *Biochim Biophys Acta* 1697, 259-269.

Taylor SS, Kim C, Vigil D, Haste NM, Yang J, Wu J & Anand GS. (2005) *Biochim Biophys Acta* 1754: 25-37.

Taylor SS, Haste NM & Ghosh G. (2005) *Cell* 122: 823-825.

Tchieu JH, Fana F, Fink JL, Harper J, Nair TM, Niedner RH, Smith DW, Steube K, Tam TM, Veretnik S, *et al*. (2003) *Nucleic Acids Res* 31: 342-344.

Tonelli, M.; Masterson, L.R.; Hallenga, K; Veglia, G.; Markely, J.L. (2007) *J. Biomol. NMR* 39, 177-185.

Traaseth, N. J.; Verardi, R.; Veglia, G. (2008) *J. Am. Chem. Soc. 130*, 2400-2401.

Traaseth, N.J., Ha, K.N., Verardi, R., Shi, L., Buffy, J.J., Masterson, L.R., Veglia, G. (2008) *Biochemistry* 47, 3-13.

Trbovic N, *et al* (2005) *J Am Chem Soc* 127, 13504-13505.

Tugarinov V & Kay LE (2003) *J Mol Biol* 327, 1121-33.

Tugarinov V, Hwang PM & Kay LE (2004) *Annu Rev Biochem* 73, 107-46.

Tzakos AG, Grace CR, Lukavsky PJ & Riek R (2006) *Annu Rev Biophys Biomol Struct* 35, 319-342.

Velyvis A, Yang YR, Schachman HK & Kay LE. (2007) *Proc Natl Acad Sci U S A* 104: 8815-8820.

Vigil D, Blumenthal DK, Taylor SS & Trewhella J. (2006) *J Mol Biol* 357: 880-889.

Vogtherr M, *et al* (2005) *J Biomol NMR* 32, 175.

Vogtherr M, *et al* (2006) *Angew Chem Int Ed Engl* 45, 993-997.

Volkman BF, Lipson D, Wemmer DE & Kern D. (2001) *Science* 291: 2429-33.

Vuister GW, Wang AC & Bax A (1993) *J Am Chem Soc* 115, 5334-5335.

Wagner G (1993) *J Biomol NMR* 3, 375-385.

Wirmer, J.; Schwalbe, H.; (2002) *J. Biomol. NMR* 23, 47-55.

Wolfram Research I. (2005) *Mathematica.*

Xu, R.; Ayers, B; Cowburn, D.; Muir, T.W.; (1999) *Proc. Natl. Acad. Sci. USA* 96, 388-393.

Yabuki, T.; Kigawa, T.; Dohmae, N.; Takio, Kl; Terada, T.; Ito, Y.; Laue, E.D.; Cooper, J.A.; Kainosho, M.; Yokoyama, S.; (1998) *J. Biomol. NMR* 11, 295-306.

Yang J, Ten Eyck LF, Xuong NH & Taylor SS. (2004) *J Mol Biol* 336: 473-487.

Yang J, Garrod SM, Deal MS, Anand GS, Woods J,Virgil L. and Taylor S (2005) *J Mol Biol* 346, 191-201.

Yonemoto WM, McGlone ML, Slice LW & Taylor SS (1991) *Methods Enzymol* 200, 581-596.

Zamoon, J., Mascioni, A., Thomas, D.D., Veglia, G., (2003) *Biophys. J.* 85, 2589-2598.

Zhang J, Ma Y, Taylor SS & Tsien RY. (2001) *Proc Natl Acad Sci U S A* 98: 14997-15002.

Appendix I

NMR, Mass Spectrometry and Chemical Evidence Reveal a Different Chemical Structure for Methanobactin that Contains Oxazolone Rings.

Lee A. Behling, Scott C. Hartsel, David E. Lewis, Alan A. DiSpirito, Dong W. Choi, Larry R. Masterson, Gianluigi Veglia, and Warren Gallagher

Reprinted with permission from the Journal of the American Chemical Society (2008) *130, 12604-12605.*

Methanobactin (mb) is a small copper-binding peptide produced by methanotrophic bacteria and is intimately involved in both their copper metabolism and their role in the global carbon cycle. The structure for methanobactin comprises seven amino acids plus two chromophoric residues that appear unique to methanobactin. In a previously published structure, both chromophoric residues contain a thiocarbonyl attached to a hydroxyimidazolate ring. In addition, one is attached to a pyrrolidine ring, while the other to an isopropyl ester. A published X-ray determined structure for methanobactin shows these two chromophoric groups forming an N_2S_2 binding site for a single Cu(I) ion with distorted tetrahedral geometry. In this report we show that NMR, mass spectrometry, and chemical data, reveal a chemical structure that is significantly different than the previously published one. Specifically, the ^{1}H and ^{13}C NMR assignments are inconsistent with an N-terminal isopropyl ester and point instead to a 3-methylbutanoyl group. Our data also indicate that oxazolone rings instead of hydroxyimidazolate rings form the core of the two chromophoric residues. Because these rings are directly involved in the binding of Cu(I) and other metals by methanobactin, and are likely involved in the many chemical activities displayed by methanobactin, their correct identity is central to developing an accurate and detailed understanding of methanobactin's many chemical and biological roles. For example, the oxazolone rings make methanobactin structurally more similar to other bacterially produced bactins and siderophores and suggest pathways for its biosynthesis.

Introduction

Methanobactin (mb) is a small, copper binding peptide produced by methanotrophic bacteria, which use methane as their primary source of energy and play an integral role in the global carbon cycle. Mb can be isolated from the growth media of methanotrophs (1-3) and is also found within the bacterial cells and associated with the copper and iron containing particulate methane mono-oxygenase (pMMO) enzyme (4, 5). pMMO, along with its soluble form (sMMO) produced under limiting copper levels, is responsible for catalyzing the oxidation of methane to methanol. Thus, this enzyme has received much attention for its role in removing a potent greenhouse gas from the atmosphere, and for the development of new hydroxylation catalysts that could make more efficient use of natural gas (6).

Mb is involved in many biological processes, including scavenging copper from the environment (1, 7, 8), serving as a copper chaperone for pMMO (1, 4, 7), complexing reactive oxygen species (4, 9), and mediating both electron flow to pMMO (3, 4, 9) and the genetic expression of pMMO (1, 10). Copper-free mb can be isolated under low copper levels (3). It binds Cu(II) with subnanomolar affinity (8), reducing it to Cu(I) (3, 11, 12). In addition, mb can bind Cd(II), Co(II), Fe(III), Mn(II), Ni(II), Zn(II), and Pb(II), and can bind and reduce Ag(I), Au(III), and possible Hg(II) (14).

The best characterized mb peptide is isolated from *Methylosinus trichosporium* OB3b (2, 11). It contains seven amino acids and two unusual chromophoric residues, each containing a thiocarbonyl/enethiol attached to a conjugated ring system. Through a combination of mass spectrometry and X-ray crystallography, these ring systems were

assigned as hydroxyimidazolates (Figure S1), which were ligated to Cu(I) in an N_2S_2 distorted tetrahedral geometry (2, 11, 13). However, it has been difficult to reconcile the chemical composition of these residues with possible biosynthetic pathways that bacteria use to produce mb. Since these residues directly participate in metal ion binding and are likely to be involved in many other chemical activities displayed by mb (9), it is crucial to establish their correct chemical structures.

In this communication, we report a revised structure for mb, which reconciles chemical analysis, NMR, X-ray, and mass spectrometry data with possible biosynthetic pathways for expression of mb in bacteria. The revised structure also clears up confusion in the literature (2, 7, 11, 12) on reporting both the chemical structure and the chemical formula for mb.

Materials and Methods

We isolated mb as described by Choi *et al.* (3) and incubated it with between 0.5 and 1.0 Cu(II) ions per mb to obtain the Cu(I)-bound form. We purified the Cu(I)-bound mb by HPLC, and after lyophilization, re-dissolved the peptide in either 100% D_2O or 90%H_2O/10% D_2O containing 9 mM sodium phosphate (pH 6.5) for NMR analysis (see Supporting Information). We collected the NMR spectra at 400 or 600 MHz at either 5 or 25 °C. We used a combination of [^{1}H,^{1}H]-COSY, [^{1}H,^{1}H]-TOCSY, [^{1}H,^{1}H]-ROESY, [^{1}H,^{15}N]-HSQC, [^{1}H,^{13}C]-HSQC, and [^{1}H,^{13}C]-HMBC experiments to assign all of the nonhydroxyl ^{1}H and ^{15}N resonances, along with all of the ^{13}C resonances except for the ones belonging to the conjugated rings (Table S1). While the sequential ^{1}H and ^{13}C assignments are consistent with the seven amino acid residues in the published structure

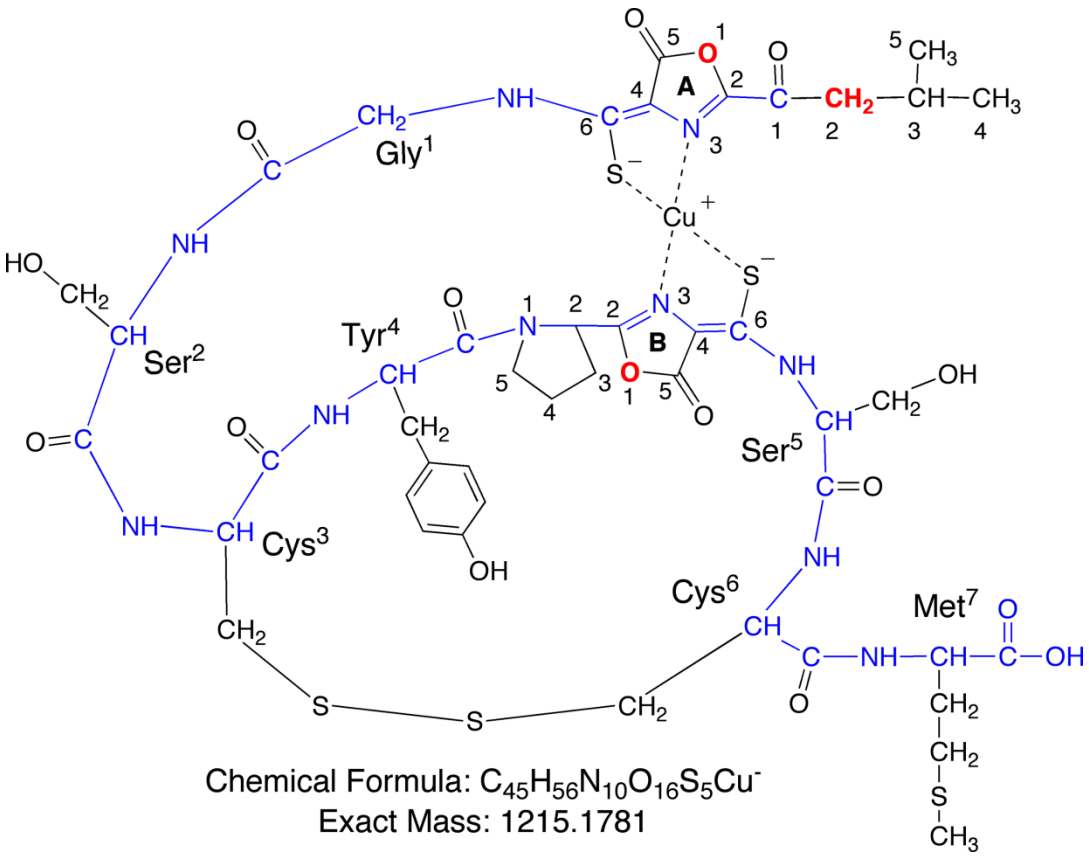


Figure 1: Revised structure for methanobactin; 1-(*N*-[mercapto-{5-oxo-2-(3-methylbutanoyl)oxazol-(*Z*)-4-ylidene}methyl]-Gly[1]-L-Ser[2]-L-Cys[3]-L-Tyr[4])-pyrrolidin-2-yl-(mercapto-[5-oxo-oxazol-(*Z*)-4-ylidene]methyl)-L-Ser[5]-L-Cys[6]-L-Met[7]. The model shown is the $(M\text{-}2H^{+}{}^{63}Cu)^{1-}$ charged species, with the backbone traced in blue. The published X-ray crystal structure (11) shows the chiral amino acids are all the L-enantiomers, the configurations of the exocyclic double bonds are both *Z*, and the chiral carbon 2 of the pyrrolidine ring is *S*.

(Figure S1), those for the N-terminal chromophoric residue are not. As shown in Figure 2, the spin system observed in [^{1}H,^{1}H]-TOCSY spectra for the N-terminal alkyl group is consistent with an isobutyl group rather than the previously reported isopropyl group (11). The [^{1}H,^{13}C]-HMBC experiment confirmed that this isobutyl moiety is attached to a carbonyl to form a 3-methylbutanoyl group.

Results

For a full characterization of the nitrogen content of the peptide, we used ^{15}N direct detection on a uniformly ^{15}N-labeled sample of Cu(I)-bound mb. Figure 3 shows the 1D proton decoupled ^{15}N spectrum. We observed only 10 ^{15}N resonances with nearly equal intensities. This confutes the previously published structure with hydroxyimidazolates rings, which should display a total of 12 nitrogen resonances. Eight nitrogen resonances are attached to protons and were assigned using an [^{1}H,^{15}N]-HSQC experiment (Figure S2). The remaining two nitrogen resonances at 226 and 249 ppm show no proton coupling and require long cycle times (>10 s) to be observed, suggesting long longitudinal spin relaxation times (T_1). Guided by the X-ray structure, we assigned these resonances to nitrogens belonging to two oxazolone rings and not hydroxyimidazolate rings.

To confirm the mass of the peptide, we carried out ESI-TOF mass spectrometry on Cu(I)-bound mb. In agreement with previously published data (2, 11, 12),we observe an *m/z* of 1215 for the (M - 2H + ^{63}Cu)$^{1-}$ species. The presence of a 3-methylbutanoyl group instead of an isopropyl ester should result in a theoretical *m/z* of 1213. However, the inclusion of the two alkylidene oxazolone rings (instead of the originally reported

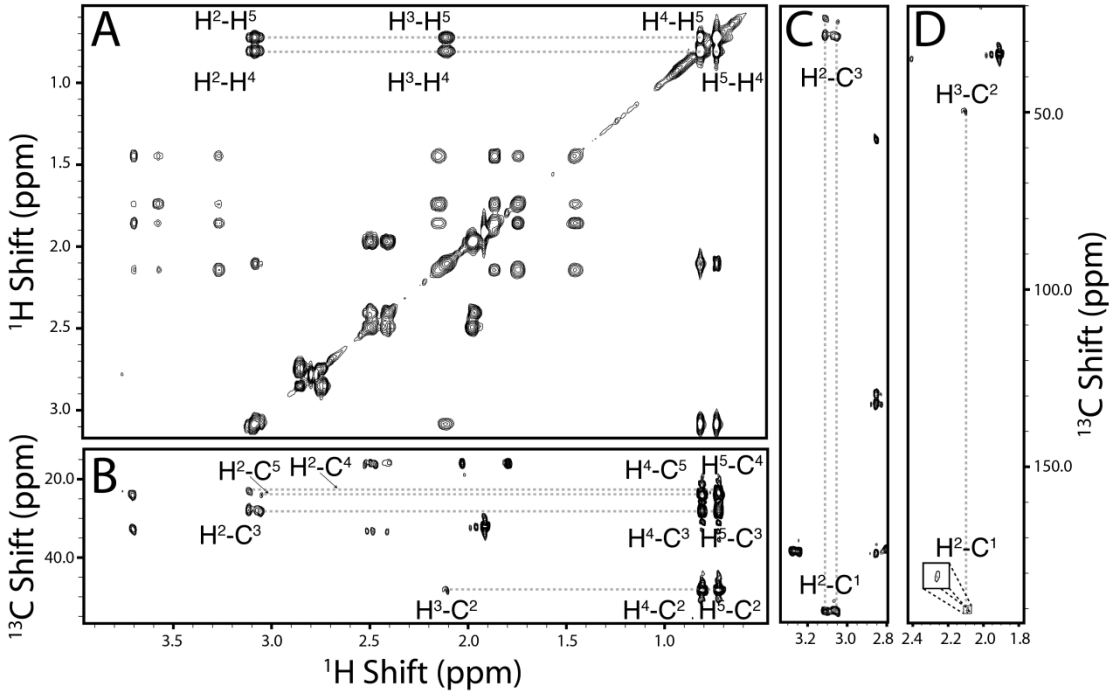


Figure 2. 600 MHz spectra of methanobactin in 9 mM phosphate/10% D_2O, pH6.5, 25°C. The sample was exposed to Cu(II) at 0.55 Cu:mb prior to isolation by HPLC. Selected regions of NMR spectra are shown indicating the presence of a 3-methylbutanoyl group. Connectivities and chemical shifts from A) [^{1}H,^{1}H]-TOCSY and B) [^{1}H,^{13}C]-HMBC spectra indicate an A_3B_3MXY type spin system (ie, $CH_3(CH_3)CHCH_2$–). Additional multiple bond connectivity from [^{1}H,^{13}C]-HMBC spectra in C) and D) confirm attachment of a carbonyl group to this moiety.

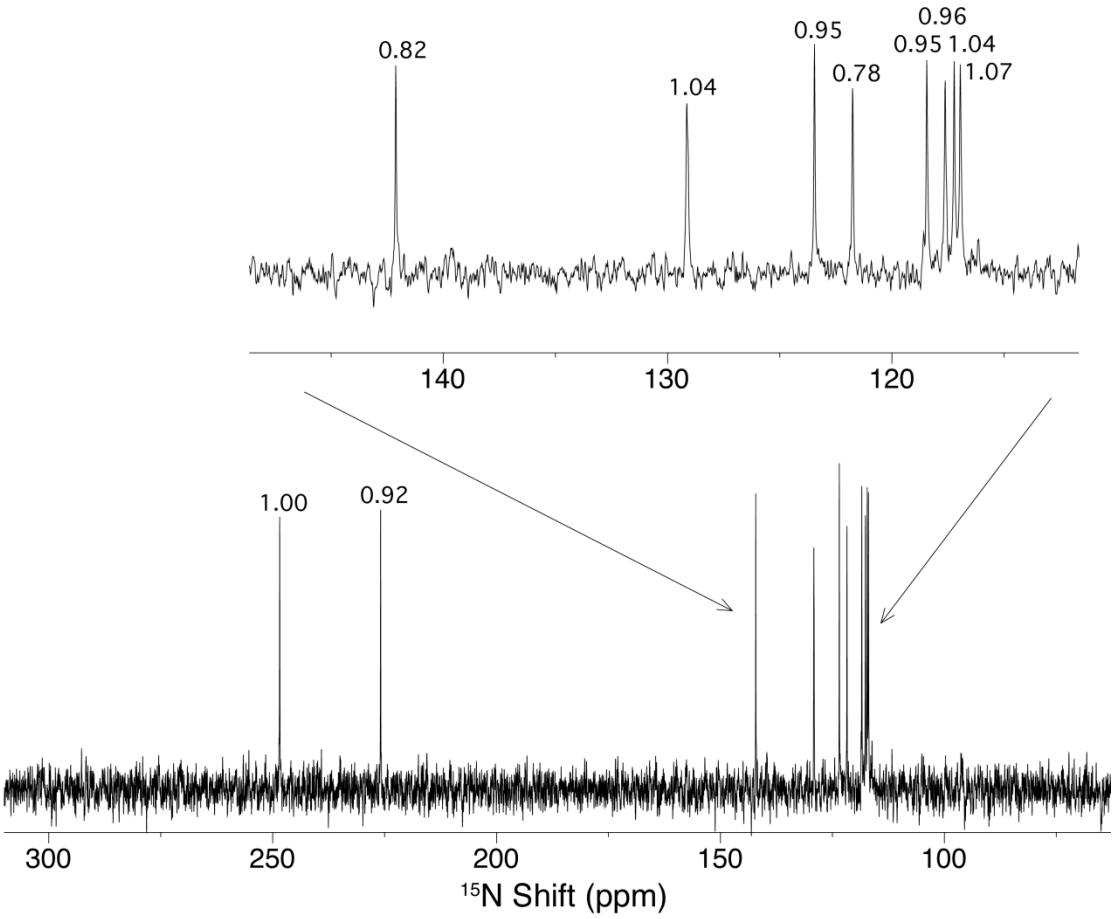


Figure 3: Proton-decoupled, 15N spectrum of U-15N Cu(I)-bound methanobactin at 25°C. A total of 10 resonances are observed with the relative integrated intensities for each indicated.

hydroxyimidazolate rings) results in an expected *m/z* of 1215, which agrees well with our measurements. We repeated the mass spectrometry analysis multiple times, obtaining values for *m/z* that are within ±1.2 ppm of that expected for our revised structure (1215.1781), but lying 8 to 10 ppm below that expected for the previously reported structure (11) (1215.1893) (Figure S3a).

The presence of oxazolone rings was also substantiated by chemical degradation of uncomplexed mb using methanolysis. Oxazolone rings contain a lactone that is susceptible to methanolysis (15) and the concomitant loss of the conjugated ring systems can be followed by UV/Vis spectrophotometry (8). Interestingly, the oxazolone B ring is more susceptible to methanolysis than the A ring (Figure S4) (8). As expected for the methanolysis of oxazolone rings, the mass of the product formed upon methanolysis of the oxazolone B ring corresponds to an increase of one equivalent of methanol (Figure S5). The opening of the oxazolone A ring requires higher concentrations of methanolic-HCl (0.01 *vs.* 0.001 saturated methanolic-HCl) and results in an additional increase in the mass by a second equivalent of methanol.

In summary, NMR, mass spectrometry and chemical degradation all point to a primary structure for mb containing two alkylidene oxazolone rings rather than the previously reported hydroxyimidazolate rings. This new revised mb structure also suggests possible pathways for the biosynthesis of mb from common amino acids. A scheme illustrating how mb could be synthesized by modifications of a peptide (LSGSCYPSSCM) is reported in the Supporting Information (Scheme S1). There are two proposed pathways. One is consistent with the fact that oxazolones have long been used as intermediates in the chemical syntheses of amino acids.[15] The other is consistent

with numerous examples of bacterially produced bactin and siderophore peptides containing the related oxazoline ring (16).

References

(1) DiSpirito, A. A.; Zahn, J. A.; Graham, D. W.; Kim, H. J.; Larive, C. K.; Derrick, T. S.; Cox, C. D.; Taylor, A. *J Bacteriol* **1998**, *180*, 3606-13.

(2) Kim, H. J.; Galeva, N.; Larive, C. K.; Alterman, M.; Graham, D. W. *Biochemistry* **2005**, *44*, 5140-8.

(3) Choi, D. W.; Antholine, W. E.; Do, Y. S.; Semrau, J. D.; Kisting, C. J.; Kunz, R. C.; Campbell, D.; Rao, V.; Hartsel, S. C.; DiSpirito, A. A. *Microbiology* **2005**, *151*, 3417-26.

(4) Choi, D. W.; Kunz, R. C.; Boyd, E. S.; Semrau, J. D.; Antholine, W. E.; Han, J. I.; Zahn, J. A.; Boyd, J. M.; de la Mora, A. M.; DiSpirito, A. A. *J Bacteriol* **2003**, *185*, 5755-64.

(5) Martinho, M.; Choi, D. W.; Dispirito, A. A.; Antholine, W. E.; Semrau, J. D.; Munck, E. *J Am Chem Soc* **2007**, *129*, 15783-5.(6) Lieberman, R. L.; Rosenzweig, A. C. *Nature* **2005**, *434*, 177-82.

(7) Balasubramanian, R.; Rosenzweig, A. C. *Curr Opin Chem Biol* **2008**, *12*, 245-249.

(8) Choi, D. W et al. *Biochemistry* **2006**, *45*, 1442-53.

(9) Choi, D. W. et al. *J Inorg Biochem* **2008**, *102*, 1571-80.

(10) Knapp, C. W.; Fowle, D. A.; Kulczycki, E.; Roberts, J. A.; Graham, D. W. *Proc Natl Acad Sci U S A* **2007**, *104*, 12040-5.

(11) Kim, H. J.; Graham, D. W.; DiSpirito, A. A.; Alterman, M. A.; Galeva, N.; Larive, C. K.; Asunskis, D.; Sherwood, P. M. *Science* **2004**, *305*, 1612-5.

(12) Hakemian, A. S.; Tinberg, C. E.; Kondapalli, K. C.; Telser, J.; Hoffman, B. M.;

Stemmler, T. L.; Rosenzweig, A. C. *J Am Chem Soc* **2005**, *127*, 17142-3.

(13) Kim, H. J. *Ph.D. Thesis, University of Kansas* **2003**.

(14) Choi, D.W. *et al*. *J Inorg Biochem* **2006**, *100*, 2150-61.

(15) Fisk, J. S.; Mosey, R. A.; Tepe, J. *J. Chem Soc Rev* **2007**, *36*, 1432-40.

(16) Roy, R. S.; Gehring, A. M.; Milne, J. C.; Belshaw, P. J.; Walsh, C. T. *Nat Prod Rep* **1999**, *16*, 249-63

Supplementary Information

Background

The published structure for methanbactin

A chemical structure for methanobactin (mb) isolated from *Methylosinus trichorsporium* Ob3b was first published by Kim *et al*, 2004 (1). This structure is shown in Figure S1 and was based on mass spectroscopy data obtained after limited hydrolyses (2, 3) of methanobactin. It also relied on an X-ray determined structure (1) of copper-bound methanobactin for the chemical identification of the two chromophoric residues. These were identified as hydroxyimidazolates and are labeled THI and HTI in Figure S1. This structure also shows the N-terminus capped with an isopropyl ester. We find the isopropyl ester is inconsistent with ^{1}H and ^{13}C NMR resonance assignments (Figure 2 and Table S1), which indicate that this group should be a 3-methylbutanoyl group (Figure 1).

Procedures for isolating copper bound methanobactin after exposure to Cu(II).

Methanobactin was isolated from the spent media of *Methylosinus trichorsporium* Ob3b as described by Choi *et al*.(4) For uniformly ^{15}N-labeled methanobactin, $K^{15}NO_3$ (Sigma Chemical Co.) was used in the culture media as the sole nitrogen source. By this procedure, methanobactin is isolated in its copper-free form, which makes it possible to carry out systematic metal ion binding studies (5,6) .These studies revealed that major

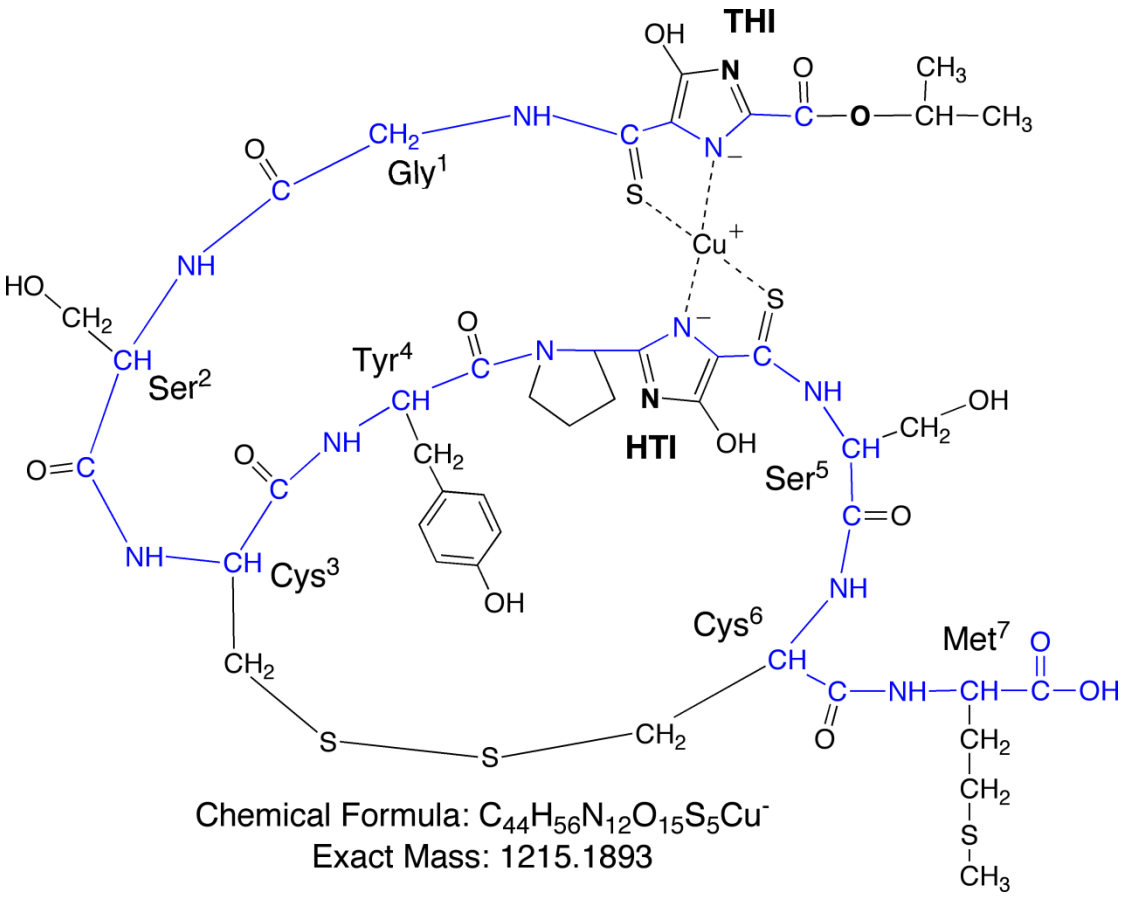


Figure S1: The previously published structure for methanobactin (1): *N*-2-Isopropylester-(4-thionyl-5-hydroxy-imidzole)-Gly[1]-Ser[2]-Cys[3]-Tyr[4]-pyrrolidine-(4-hydroxy-5-thionyl-imidazole)-Ser[5]-Cys[6]-Met[7]. The model shown has the two hydroxyimidazolate rings ligated to Cu^{1+} to form the $(M\text{-}2H+^{63}Cu)^{1-}$ charged species that is observed by mass spectrometry. The backbone is traced in blue.

spectroscopic changes associated with copper binding are complete by around 0.6 Cu per methanobactin. We have used ^{1}H 1D and [^{1}H,^{15}N]-HSQC NMR experiments to monitor Cu ion binding by methanobactin and observe a similar leveling off of changes at around 0.6 Cu:mb (data not shown). For the studies reported in this communication, we used HPLC to isolate the Cu(I)-bound methanobactin formed after exposure to between 0.5 and 1.0 Cu:mb.

To prepare Cu(I)-bound methanobactin, lyophilized, copper-free methanobactin was dissolved in a 10 mM sodium phosphate buffer at pH 6.5. The concentration of the methanobactin was determined by UV/Vis spectrophotometry using a molar extinction coefficient of $\varepsilon_{394} = 16.07\ mM^{-1}cm^{-1}$ for the unbound methanobactin and $\varepsilon_{394} = 9.75\ mM^{-1}cm^{-1}$for the Cu-bound methanobactin (6). Cu(II) was added as $CuSO_4$ in 0.1 Cu:mb increments and the *pH* readjusted to 6.5 with NaOH after each addition. After reaching between 0.5 and 1.0 Cu:mb, the sample was subjected to reverse phase HPLC on a Hamilton 305 mm x 7 mm PRP-3 column and eluted with a gradient of 1% to 99% MeOH containing 1 mM ammonium acetate buffer, *pH* 5.0. After HPLC, the fraction containing the Cu(I)-bound methanobactin was lyophilized, redissolved in water, and lyophilized a second time.

NMR experiments

For NMR analyses, the lyophilized Cu(I)-bound methanobactin was dissolved to a concentration of between 2 and 5 mM in 9 mM sodium phosphate, pH 6.5, containing 10% D_2O, or in 100% D_2O adjusted to pH 6.5. Spectra were collected at either 5 °C or

25 °C on a 400 MHz Bruker Avance II spectrometer equipped with either a z-gradient, inverse detection TXI probe, or z-gradient, broadband BBO probe. Some experiments were also carried out on a 600 MHz Varian VNMRS spectrometer equipped with a z-gradient, inverse detection HCN probe. 1D-^{1}H, [^{1}H-^{1}H]-COSY, [^{1}H-^{1}H]-TOCSY, [^{1}H-^{1}H]-ROESY, [^{1}H-^{13}C]-HSQC, [^{1}H-^{15}N]-HSQC and [^{1}H-^{13}C]- HMBC experiments were collected with the inverse detection probes. ^{15}N-1D direct-detection experiments were collected with the BBO probe. WATERGATE 3-9-19 pulse sequences were used for solvent suppression in the ^{1}H-1D, [^{1}H-^{1}H]-COSY, [^{1}H-^{1}H]-TOCSY, and [^{1}H-^{1}H]-ROESY experiments. The [^{1}H-^{1}H]-ROESY experiments were collected using a mixing time of 200 ms. [^{1}H-^{15}N]- HSQC indirect detection and ^{15}N-1D direct detection experiments were carried out on uniformly ^{15}N-labeled Cu(I)-bound methanobactin samples that were prepared as described above for the unlabeled samples. The spectra obtained from these experiments were processed using either *iNMR*[7] or *NMRPipe* (8), and analyzed using *Sparky* (9). The ^{1}H, ^{13}C and ^{15}N assignments obtained from these analyses are shown in Table S1. Figure S2 shows the [^{1}H-^{15}N]-HSQC spectrum, along with ^{1}H and ^{15}N external projections, which were used to assign the 7 amide nitrogens and the pyrrolidinyl nitrogen.

Table S1: ^{1}H, ^{13}C, and ^{15}N resonance assignments for Methanobactin from *Methylosinus trichosporus* OB3b.

		Resonances (ppm)*					Resonances (ppm)		
	Atom	^{1}H	^{13}C	^{15}N	Residue	Atom	^{1}H	^{13}C	^{15}N
3-Methyl-	C1		188.1		Pyrrolidinyl	N1			142.1
butanoyl	C2		46.3			C2		55.5	
	C3		26.0			C3		30.8	
	C4		21.6			C4		22.1	
	C5		22.2			C5		46.2	
	H2	3.07				H2	3.71		
	H2	3.11				H3	1.46		
	H3	2.11				H3	1.87		
	H4	0.82				H4	1.75		
	H5	0.73				H4	2.15		
Oxazolone A	N3			225.9 or 248.5		H5	3.28		
						H5	3.58		
Gly1	N			121.7	Oxazolone B	N3			225.9 or 248.5
	C		173.5						
	Cα		48.6		Ser5	N			117.0
	H	9.28				C		171.9	
	Hα	4.46				Cα		60.5	
	Hα	4.59				Cβ		60.7	
Ser2	N			117.6		H	8.74		
	C		171.6			Hα	4.85		
	Cα		57.5			Hβ	3.99		
	Cβ		60.6			Hβ	4.26		
	H	9.04			Cys6	N			118.4
	Hα	4.41				C		171.4	
	Hβ	3.97				Cα		57.0	
	Hβ	4.10				Cβ		40.7	
Cys3	N			117.2		H	8.12		
	C		171.0			Hα	4.81		
	Cα		52.6			Hβ	3.27		

	Cβ		44.8			Hβ	3.46		
	H	7.73			Met7	N			129.1
	Hα	5.12				C		177.9	
	Hβ	2.80				Cα		56.0	
	Hβ	3.78				Cβ		31.4	
Tyr4	N			123.5		Cγ		30.1	
	C		171.9			Cε		14.1	
	Cα		55.1			H	7.80		
	Cβ		36.2			Hα	4.12		
	Cγ		126.3			Hβ	1.98		
	Cδ		130.1			Hβ	1.98		
	Cε		115.3			Hγ	2.41		
	Cζ		155.2			Hγ	2.50		
	H	7.11				Hε	1.92		
	Hα	4.30							
	Hβ	2.75							
	Hβ	2.86							
	Hδ	6.57							
	Hε	6.61							

*The ^{1}H assignments were made at 25°C and referenced against TSP. The ^{13}C and ^{15}N assignments were made at 25°C.

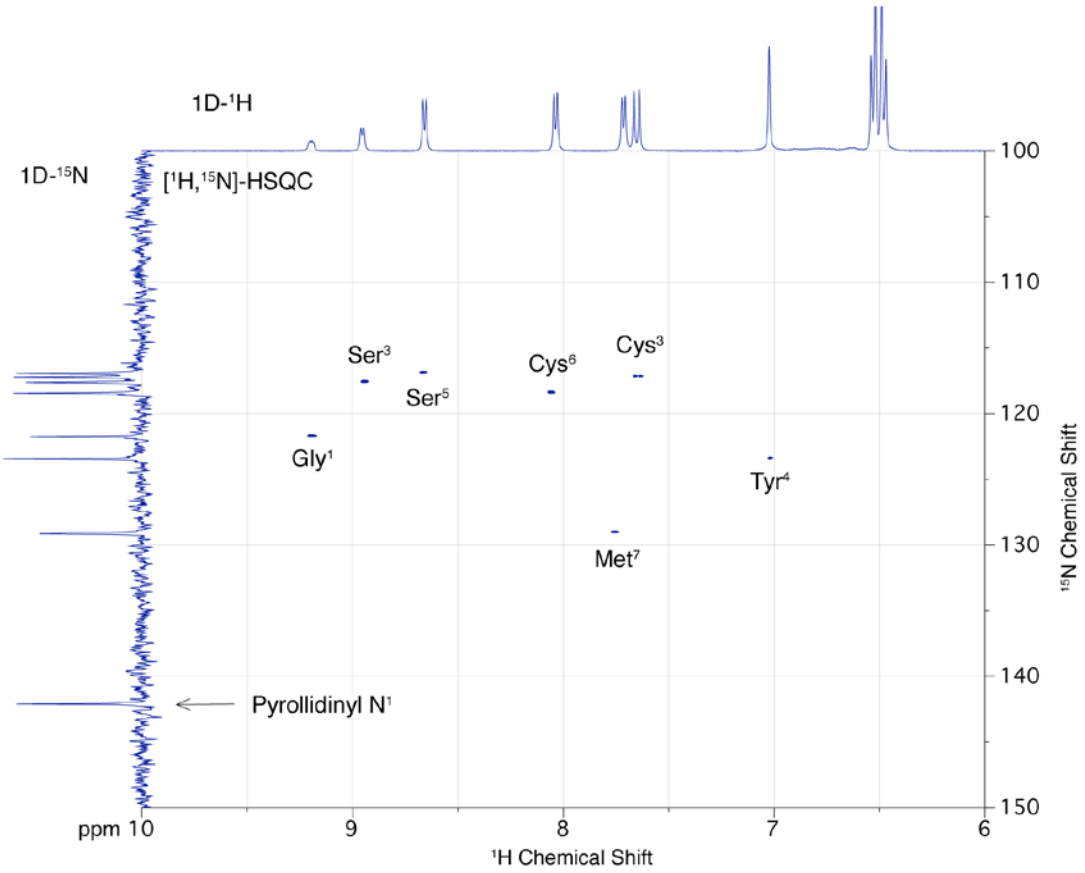


Figure S2: 400 MHz [^{1}H,^{15}N]-HSQC spectrum of uniformly 15N-labeled methanobactin in 9 mM phosphate/10% D_2O, *pH*6.5, 25°C. The sample was exposed to Cu(II) at 0.50 Cu:mb prior to isolation by HPLC. The ^{15}N projection was collected separately in a 1D-^{15}N direct detection experiment on the same sample. The 1D-^{1}H projection was collected separately on an unlabeled sample of methanobactin.

ESI-TOF mass spectroscopy

Accurate mass experiments were carried out using an Agilent 6210 ESI-TOF mass spectrometer operating in negative ion mode. Samples dissolved in 70% acetonitrile were injected directly into the electrospray source of the mass spectrometer. The sample used to collect the data shown in Figure S3b also contained a mixture of 10 reference masses with *m/z's* of between 300 and 2800. The solution containing these references masses was supplied as a calibrating solution for the spectrometer (Agilent ESI-TOF Tuning Mix). The specification of accuracy for this instrument, when using internal reference masses, is 2 ppm. Before running a methanobactin sample, the mass spectrometer was tuned and calibrated using the ES TOF Tuning Mix. After collecting a mass spectrum, the spectra were re-calibrated using 7 of the internal reference masses. These had *m/z's* ranging from 302 to 2234. Figure S3a shows the *m/z* values obtained from four trials and compares them to the *m/z* values expected for the previously published and the proposed structures for methanobactin. Figure S3b shows a representative example of a mass spectrum for Cu(I)-bound methanobactin and compares it to the isotropic distribution expected for molecule with the proposed chemical formula shown in Figure 1.

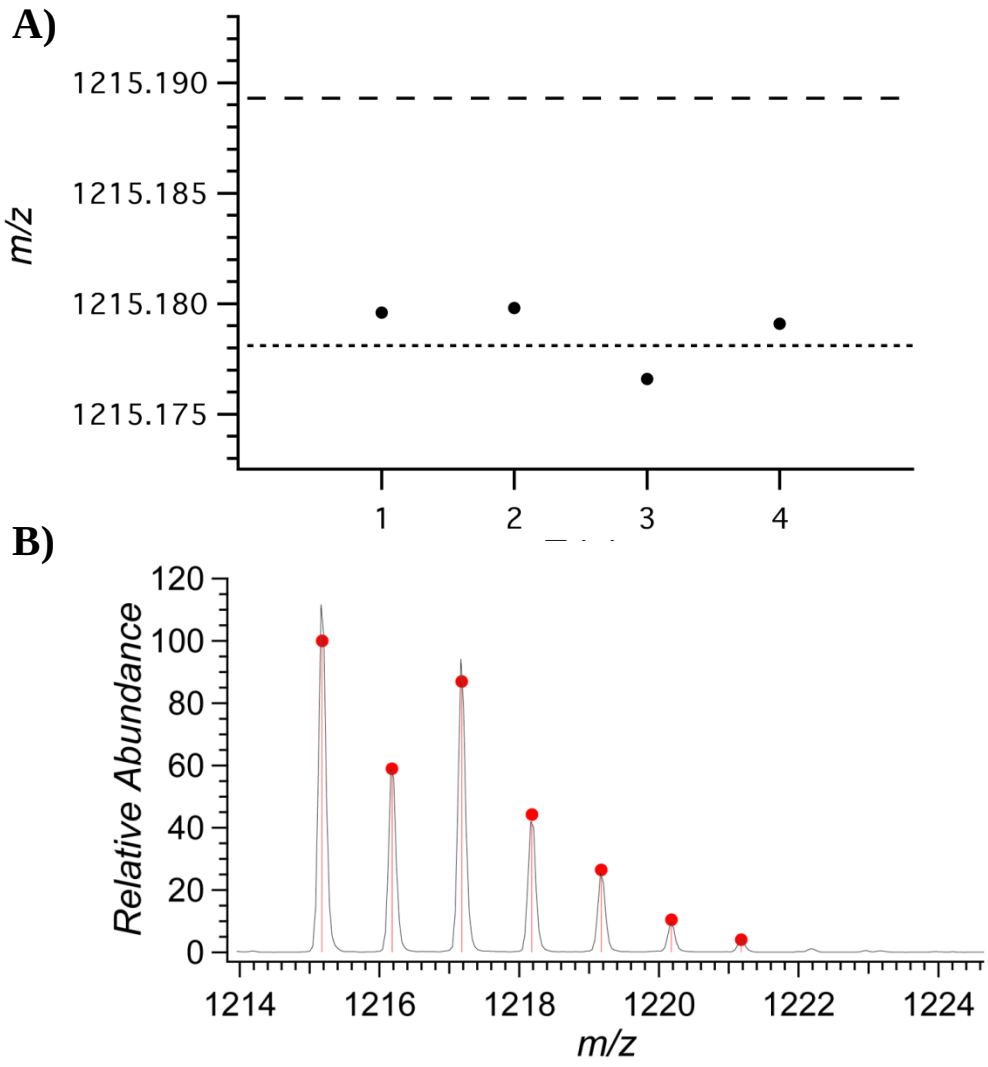


Figure S3: a) The accurate mass determination of (M-2H+63Cu)1- species of Cu(I)-bound methanobactin. The results of four trials compared to the expected mass of the published structure (dashed line) and the proposed structure (dotted line). **b)** The mass spectrum for trial 1 shown in a). The red dots and lines indicate the expected isotope pattern for a molecule with the same molecular formula.

Methanolysis of uncomplexed methanobactin

Methanolysis of uncomplexed methanobactin was carried out by dissolving lyophilized methanobactin in methanolic HCl. Saturated Methanolic HCl was produced by bubbling HCl gas through methanol for 5 minutes. The saturated methanolic HCl was then diluted to 1/10, 1/100 and 1/1000 with MeOH. 1 mg samples of methanobactin, were dissolved in 1 mL MeOH, and 1 mL aliquots of the diluted and undiluted saturated methanolic HCl. The mixtures were allowed to react at room temperature for 15 minutes and then rotary-evaporated to a dry film at 30°C. The evaporation took approximately 3-4 minutes. The dried films were taken up in 1 mL H_2O and the *pH* adjusted to between 6.7 and 6.9 with 100 mM NaOH. The samples were then diluted to 1/10 with water and analyzed by UV/Vis spectrophotometry and ESI-TOF mass spectroscopy. Figure S4 shows an overlay of the UV/Vis spectra from these samples. The spectrum for the sample that contained no HCl is characteristic of uncomplexed methanobactin. As the HCl concentration is increased it leads first to the loss of the oxazolone B ring, which is associated with the absorbance at 340 nm (5), followed with the loss at higher methanolic HCl concentrations of the oxazolone A ring, which is associated with the absorbance at 394 nm (5).

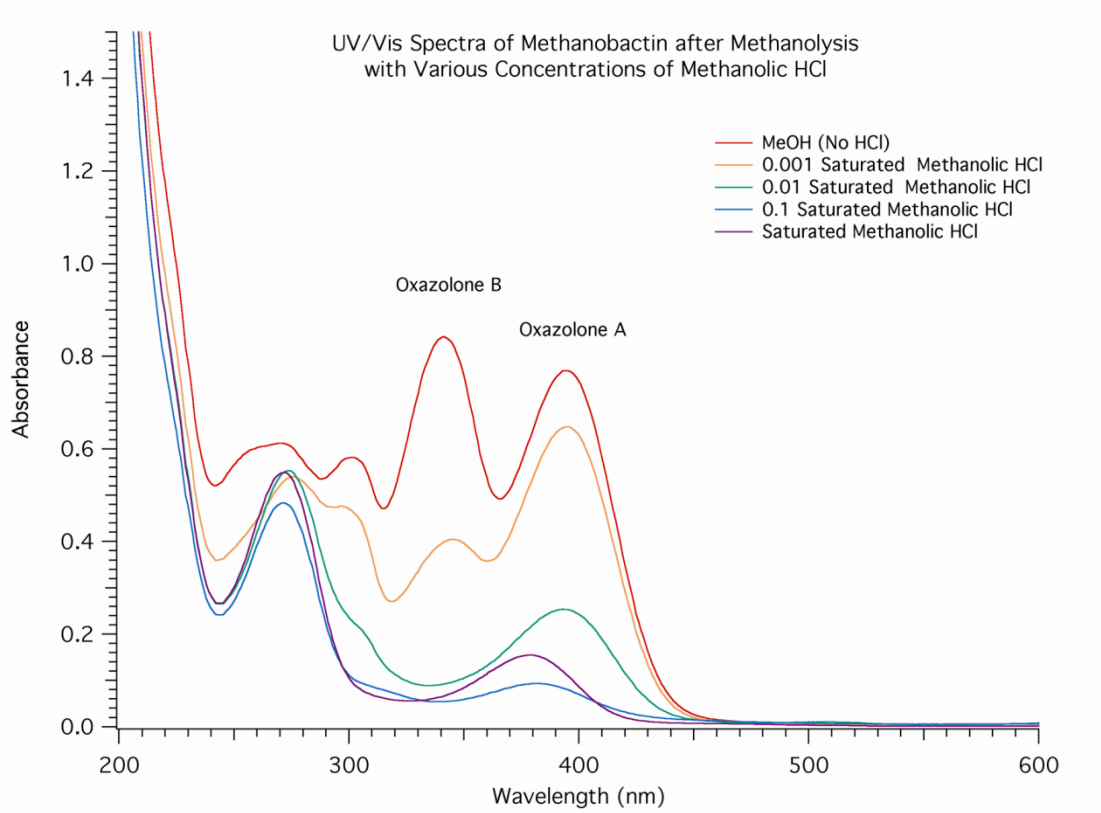


Figure S4: The UV/Vis spectra of methanobactin after methanolysis with various concentrations of methanolic HCl. The assignment of the the A_{394} to the N-terminal chromophoric residue and A_{340} to the other chromophoric residue was made based on the more extended conjugated system for the N-terminal residue (5).

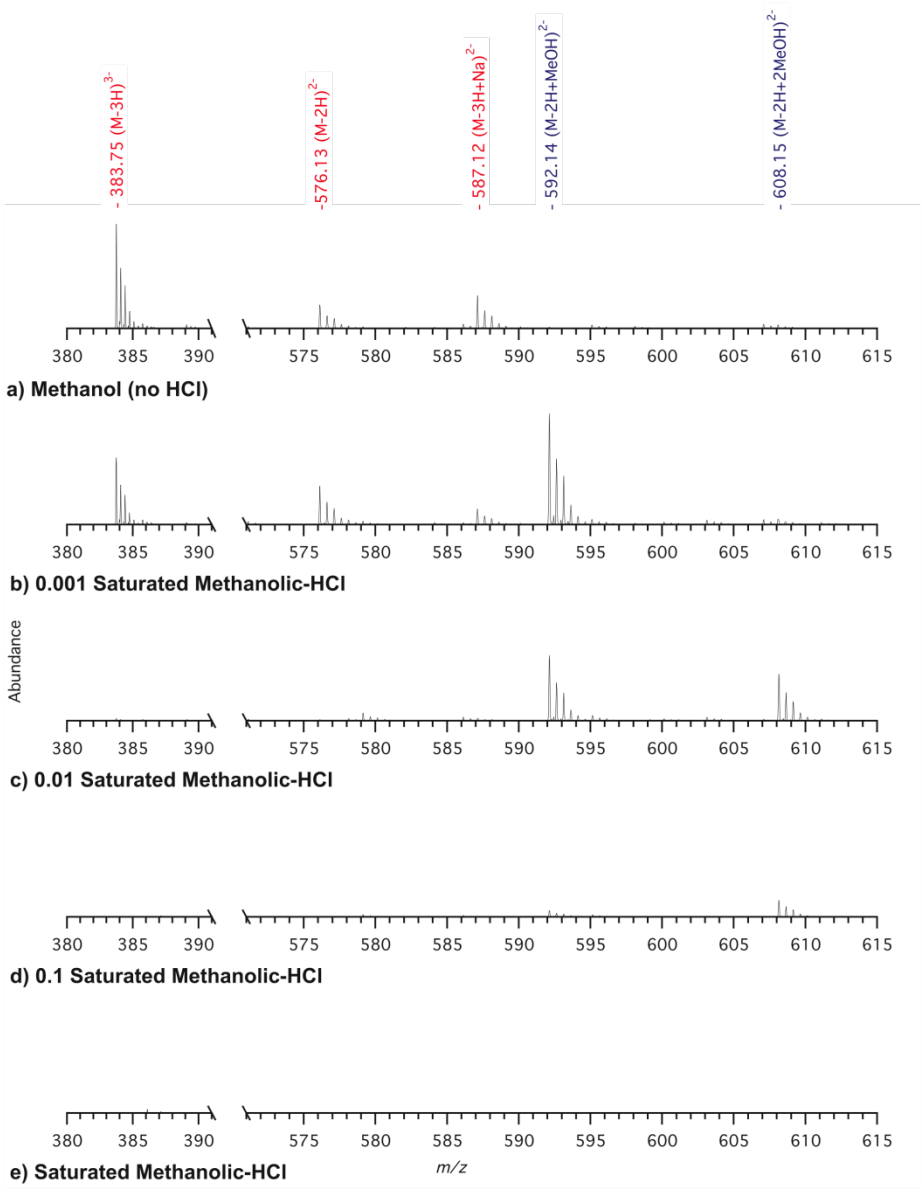


Figure S5: The ESI-TOF mass spectra for the same samples shown in Figure S4. The observed *m/z* values for the methanobactin species, $[(M-3H)^{3-}$, $(M-2H)^{2=}$, and $(M-3H+Na)^{2-}]$ are labeled in red, while those for the products of the methanolysis reaction, $[(M-2H+MeOH)^{2-}$ and $(M-2H+2MeOH)^{2-}]$ are labeled in blue. In the 0.1 saturated methanolic-HCl and the saturated methanolic-HCl reactions the methanobactin species were replaced by species with lower masses (not shown).

Figure S5 shows the corresponding ESI-TOF mass spectra for the samples used in obtaining the UV/Vis spectra shown in Figure S5. The sample that contained no HCl showed species that are all associated with uncomplexed methanobactin, $(M-3H)^{3-}$, $(M-2H)^{2-}$ and $(M-3H+Na)^{2-}$. As the HCl concentration was increased, these species were replaced by two new species with higher masses, first one with a charge of 2- at *m/z* 592.14, followed by one with a charge of 2- at *m/z* 608.15. Figure S6 shows that these two species correspond to what is expected for methanobactin which contains oxazolone rings that have undergone methanolysis (10), first at one of its oxazolone rings, $(M-2H+MeOH)^{2-}$, followed by both rings $(M-2H+2MeOH)^{2-}$. At even higher concentrations of methanolic HCl, species with lower masses replaced these species, indicating further acid degradation of the methanobactin.

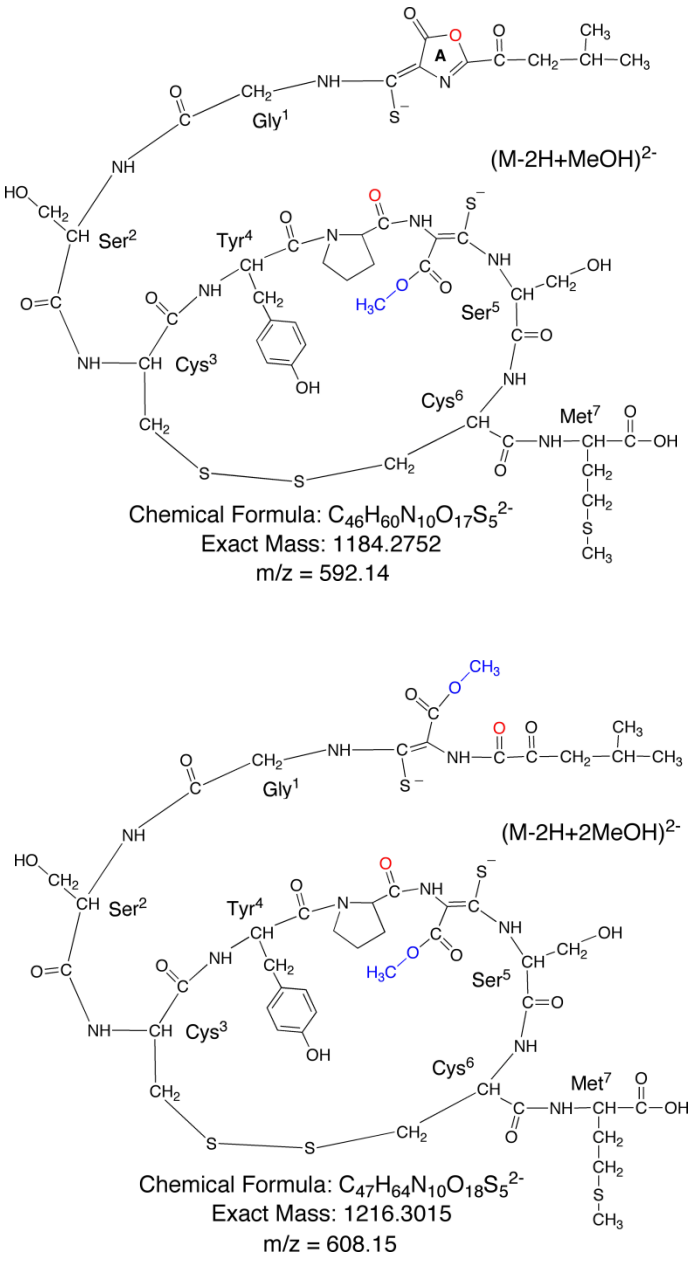


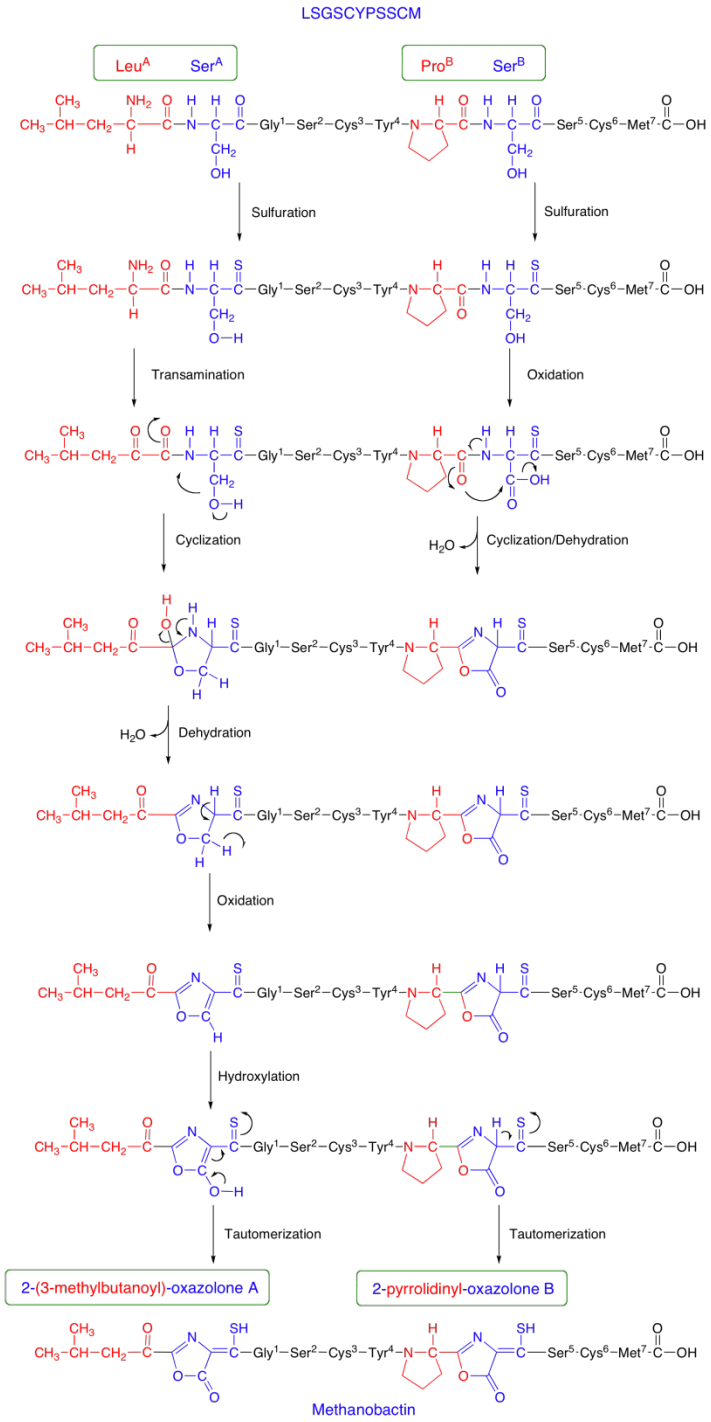
Figure S6: The chemical structures and m/z values for the expected products from the methanolysis of methanobactin containing two oxazolone rings.

Suggested pathways for the biosynthesis of methanobactin

The chemical structure that we propose for methanobactin, which contains alkylidene oxazolone rings, suggests a couple of pathways for the biosynthesis of methanobactin from an oligopeptide. Oxazoline rings, which are related to oxazolone rings, are commonly found in other bactins and siderophores (11). These molecules are synthesized by non-ribosomal peptide synthetases and incorporate oxazoline rings into the backbone of an oligopeptide by reacting the sidechain hydroxyl group of a serine with the carbonyl carbon of the preceding residue in the sequence. The hydroxylation and oxidation of the oxazoline ring in the proposed biosynthesis of methanobactin would require some modifications to this pathway to incorporate the oxo substituent on the ring. An example of how this strategy could lead to the synthesis of the 3-methylbutanoyl-oxazolone A ring is shown on the left-hand side of Scheme S1.

Another possible pathway is suggested by a side reaction that is observed in solid-state peptide synthesis and that leads to the incorporation of an oxazolone ring at the C-terminus of a peptide (12). The reaction involves reacting an amide carbonyl oxygen with the the carbonyl carbon of the next residue in the sequence. On the right-hand side of Scheme S1 we show how a modification of this reaction, to involve the side chain carbonyl formed from the oxidation of a serine side chain, could lead to the incorporation of oxazolone B ring into methanobactin. The incorporation of the 3-methylbutanoyl group at the N-terminus could arise from a transamination of a leucine residue, while no further modification is required to incorporate the pyrrolidine ring from a proline residue. Sulfuration of the two amide carbonyls attached to the two

oxazolone rings would produce thiocarbonyls, which could then tautomerize to the enethiols shown in the proposed structure.



Scheme S1: Proposed pathways for the biosynthesis of methanobactin. The proposed structure shown in Figure 1 suggests a couple of pathways for the synthesis methanobactin from the 11 residue oligopeptide LSGSCYPSSCM. Using a

modification of the reactions that lead to the biosynthesis oxazoline rings in bactins and siderophores (11), Leu^{A} and Ser^{A} are shown combining to form the 3-methylbutanoyl group and the oxazolone A ring. And using a modification of a side reaction that is observed in solid-state peptide synthesis (12), Pro^{B} and Ser^{B} are shown combining to form the pyrollidinyl group and the oxazolone B ring.

References

(1) Kim, H. J.; Graham, D. W.; DiSpirito, A. A.; Alterman, M. A.; Galeva, N.; Larive, C. K.; Asunskis, D.; Sherwood, P. M. *Science* **2004**, *305*, 1612-5.

(2) Kim, H. J.; Galeva, N.; Larive, C. K.; Alterman, M.; Graham, D. W. *Biochemistry* **2005**, *44*, 5140-8.

(3) Kim, H. J. *Ph.D. Thesis, University of Kansas* **2003**.

(4) Choi, D. W.; Antholine, W. E.; Do, Y. S.; Semrau, J. D.; Kisting, C. J.; Kunz, R. C.; Campbell, D.; Rao, V.; Hartsel, S. C.; DiSpirito, A. A. *Microbiology* **2005**, *151*, 3417-26.

(5) Choi, D. W.; Zea, C. J.; Do, Y. S.; Semrau, J. D.; Antholine, W. E.; Hargrove, M. S.; Pohl, N. L.; Boyd, E. S.; Geesey, G. G.; Hartsel, S. C.; Shafe, P. H.; McEllistrem, M. T.; Kisting, C. J.; Campbell, D.; Rao, V.; de la Mora, A. M.; Dispirito, A. A. *Biochemistry* **2006**, *45*, 1442-53.

(6) Choi, D. W.; Do, Y. S.; Zea, C. J.; McEllistrem, M. T.; Lee, S. W.; Semrau, J. D.; Pohl, N. L.; Kisting, C. J.; Scardino, L. L.; Hartsel, S. C.; Boyd, E. S.; Geesey, G. G.; Riedel, T. P.; Shafe, P. H.; Kranski, K. A.; Tritsch, J. R.; Antholine, W. E.; DiSpirito, A. A. *J Inorg Biochem* **2006**, *100*, 2150-61.

(7) Giuseppe Balacco, *iNMR*, Nucleomatica, Via Francesco Carabellese, 10/C,

70056 Molfetta BA, Italy (http://www.inmr.net/)

(8) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J Biomol NMR* **1995**, *6*, 277-93.

(9) T. D. Goddard and D. G. Kneller, *SPARKY 3*, University of California, San Francisco (http://www.cgl.ucsf.edu/home/sparky/)

(10) Fisk, J. S.; Mosey, R. A.; Tepe, J. *J. Chem Soc Rev* **2007**, *36*, 1432-40.

(11) Roy, R. S.; Gehring, A. M.; Milne, J. C.; Belshaw, P. J.; Walsh, C. T. *Nat Prod Rep* **1999**, *16*, 249-63.

(12) Benoiton, N. L. *Chemistry of Peptide Synthesis*; CRC Press: Boca Raton, FL, **2006**.

(13) Choi, D. W.; Semrau, J. D.; Antholine, W. E.; Hartsel, S. C.; Anderson, R. C.; Carey, J. N.; Dreis, A. M.; Kenseth, E. M.; Renstrom, J. M.; Scardino, L. L.; Van Gorden, G. S.; Volkert, A. A.; Wingad, A. D.; Yanzer, P. J.; McEllistrem, M. T.; de la Mora, A. M.; DiSpirito, A. A. *J Inorg Biochem* **2008**, *102*, 1571-80.

Appendix II

Structural and Dynamic Basis of Phospholamban and Sarcolipin Inhibition of Ca^{2+}-ATPase (SERCA)

Nathaniel J. Traaseth, Kim N. Ha, Raffaello Verardi, Lei Shi, Jarrod J. Buffy, Larry R. Masterson, and Gianluigi Veglia

Reprinted with permission from Biochemistry (2008) *47, 3-13.*

Phospholamban (PLN) and sarcolipin (SLN) are two single pass membrane proteins that regulate Ca^{2+} ATPase (SERCA), an ATP-driven pump that translocates calcium ions into the lumen of the sarcoplasmic reticulum initiating muscle relaxation. Both proteins bind SERCA through intramembrane interactions, impeding calcium translocation. While post-translational phosphorylation of PLN at Ser-16 and/or Thr-17 reestablishes calcium flux, the regulatory mechanism of SLN remains elusive. SERCA has been crystallized in several different states along the enzymatic reaction coordinates, providing remarkable mechanistic information; however, the lack of high-resolution crystals in the presence of PLN and SLN limits the current understanding of the regulatory mechanism. This brief review offers a survey of our hybrid structural approach using solution and solid-state NMR methodologies to understand SERCA regulation from the point of view of PLN and SLN. These results have furthered the understanding of the calcium translocation process, and are the basis for designing new therapeutic approaches to ameliorate muscle malfunctions.

Introduction

Two membrane proteins, sarco(endo)plasmic reticulum calcium ATPase (SERCA) and the ryanodine receptors (Ryr), play major roles in calcium regulation within muscle cells. Ryr1 and Ryr2 are responsible for releasing Ca^{2+} from the sarcoplasmic reticulum (SR) of skeletal and cardiac muscle, respectively, resulting in muscle contraction. Likewise, SERCA1a and SERCA2a pump calcium into the SR in skeletal and heart muscle, initiating muscle relaxation. Phospholamban (PLN), a 52-residue protein spanning the SR membrane, is an endogenous inhibitor of SERCA, lowering the apparent calcium affinity of the ATPase. The relief of SERCA inhibition is achieved by phosphorylation of PLN at Ser-16 by protein kinase A and/or Thr-17 by Ca^{2+}/calmodulin-dependent protein kinase (1). *In vivo* studies demonstrate that phosphorylation at Ser-16 and Thr-17 have different roles, suggesting that these mechanisms act independently (2, 3).

Sarcolipin (SLN) has a primary sequence homologous to that of the transmembrane domain of PLN (4, 5). Initially, SLN was thought to be the counterpart of PLN within skeletal muscle, playing only an ancillary role in cardiac muscle. Recently, however, significant expression levels of SLN have been detected in cardiac atrial muscle, with lower levels in ventricular muscle (6-8), suggesting that it may play an important role in regulation of the heart . When SLN was initially co-purified with fast-twitch skeletal SERCA1a (9), no post-translational modifications were identified, which led to the conclusion that the regulation of SLN depended on its variable expression levels (10, 11). Recent evidence shows that SLN is able to regulate SERCA

and that its inhibition can be fully reversed by isoproterenol, a β-adrenergic receptor agonist, in PLN knock-out mice (12). These results led to the hypothesis that the inhibitory effect of SLN can be reversed via phosphorylation in a similar manner as PLN. *In vitro* experiments have shown that SLN can be phosphorylated at Thr-5 when co-transfected with serine/threonine kinase 16 (STK16) (12). From the biological data, it is clear that phosphorylation of PLN, and possibly SLN, constitute important driving forces for calcium re-uptake into cardiac SR.

Several crystal structures of SERCA in different conformations within the enzymatic cycle have revealed important atomic details regarding SERCA's mechanism (Figure 1) (13, 14, 15, 16, 17, 18). How much is known about the molecular details of PLN/SLN regulation and calcium translocation by SERCA? PLN is thought to bind and inhibit the low affinity calcium form of SERCA (E2) and detach from the enzyme (either partially or totally) upon phosphorylation at Ser-16, reversing its inhibitory effect and restoring the affinity of SERCA for Ca^{2+} ions. The only experimental structure of the SERCA/PLN complex is a low-resolution cryoelectron microscopy image (8-10 Å) obtained by Stokes and co-workers (19). Hampering the formation of large, highly diffracting SERCA/PLN and SERCA/SLN co-crystals is the dynamic interplay between the proteins. For this reason, MacLennan and co-workers have used a plethora of available biological data (mutagenesis studies, co-immunoprecipitation assays, and cross-linking experiments) in-concert with molecular dynamics simulations to model SERCA/PLN (20), SERCA/SLN (21), and SERCA/PLN/SLN complexes (21) (Figure 2). Hutter *et al.* have also modeled the solution structure of C41F PLN determined in chloroform/methanol with the E2 form of SERCA using molecular

mechanics (22). While these models shed light into the interaction between SERCA and PLN and SLN, there are inconsistencies concerning the topology and structure of PLN and SLN within the complexes.

This review reports on our recent progress involving the structure determination of PLN and SLN in the free forms and towards elucidation of the interaction with SERCA using solution and solid-state NMR. While many contributions to the structural analysis of PLN and SLN from other laboratories are cited and related to our work, this review is not intended to be an exhaustive overview of the large amount of structural and biological information on PLN and SLN within the literature.

Choice of Membrane Mimicking Environments for Spectroscopic Studies

For both solution and solid-state NMR, we use two major criteria for sample preparation: 1) functionality of proteins under NMR conditions and 2) the ability to acquire high quality NMR spectra. Being membrane embedded proteins, SLN and PLN need lipid environments to elicit their biological function. After scanning several different conditions, we chose dodecylphosphocholine (DPC) detergent micelles for solution NMR studies and a 4/1 mixture of Dioleoyl-*sn*-Glycero-3-Phosphocholine/Dioleoyl-*sn*-Glycero-3-Phosphoethanolamine (DOPC/DOPE) lipids for solid-state NMR studies. Figure 3A and B shows the calcium dependence of SERCA in the presence and absence of PLN (or SLN) under these conditions. The hydrolysis of ATP by SERCA is followed directly using ^{31}P NMR spectroscopy in DPC micelles and indirectly using the coupled enzyme assay in lipids (23, 24, 25). A representative

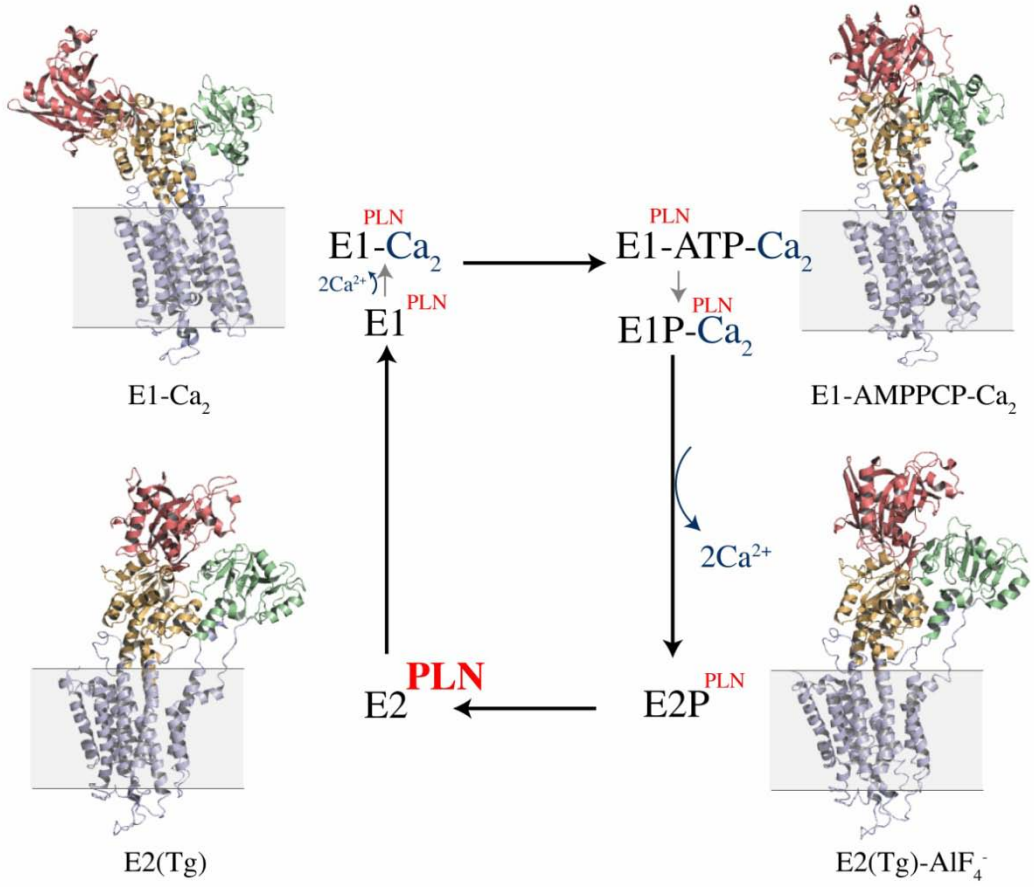


Figure 1: Enzymatic cycle mimicking the four major conformational states of SERCA: E2 (*1IWO)* (14), E1-Ca_2 (*1SU4)* (13), E1-ATP-Ca_2 (*1VFP)* (15), E2P (*1XP5)* (17). PLN and SLN are believed to inhibit the E2 conformation as indicated within the model. The model was reproduced from Inesi *et al.* (60).

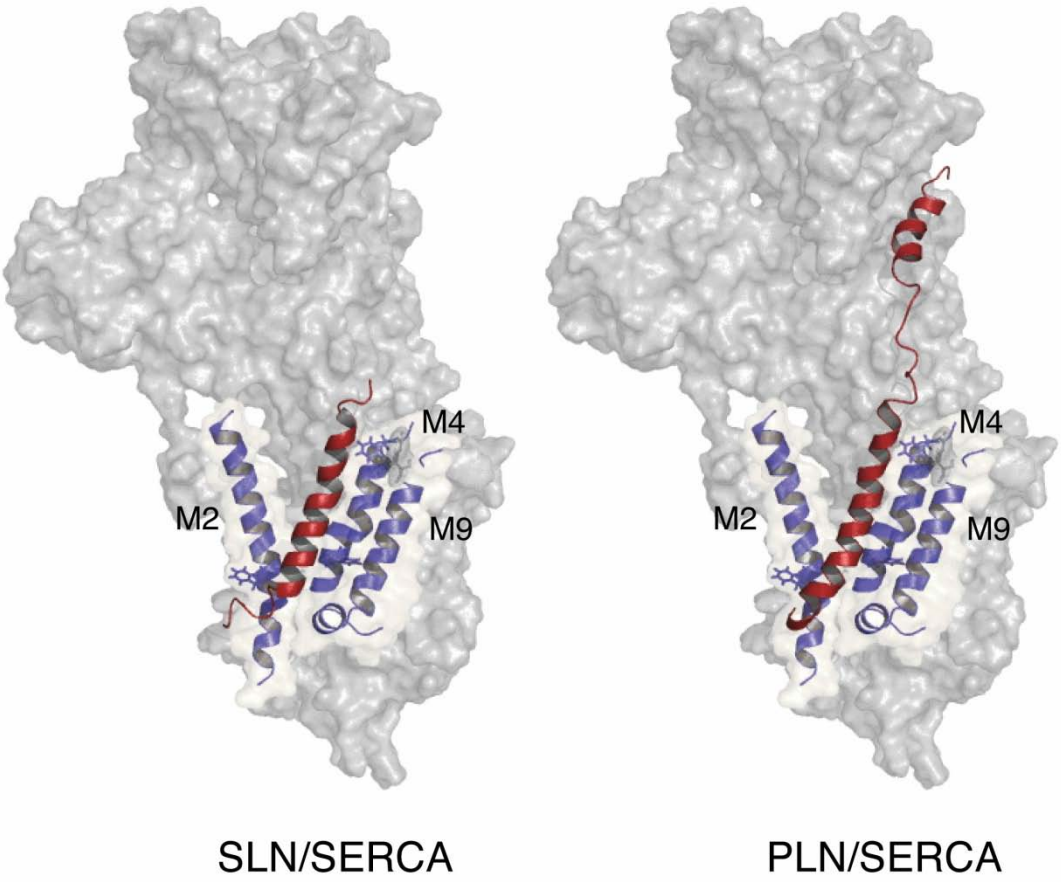


Figure 2: Models of SERCA/PLN (20) and SERCA/SLN (21). Coordinates of complexes generously provided by D.H. Maclennan.

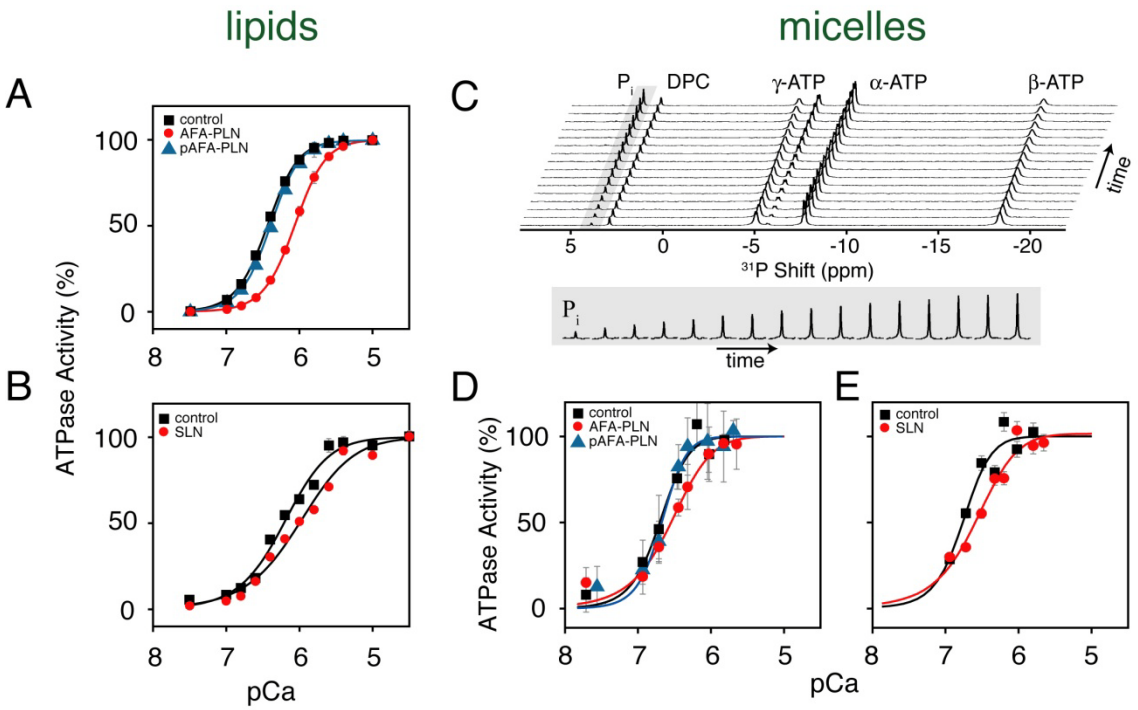


Figure 3: Activity assays of SERCA in the presence and absence of PLN and SLN in 4/1 DOPC/DOPE lipid bilayers (A and B) and DPC detergent micelles (D and E). SERCA activity in lipid bilayers was measured via the coupled enzyme assay and is reproduced from Buck *et al*. (23). The activity assays in DPC micelles were done using ^{31}P NMR spectroscopy as previously reported (25), reproduced from Buffy *et al.* (24) and Traaseth *et al*. (25).

example of the kinetic results from the ^{31}P spectroscopic assay is shown in Figure 3C. In both lipids and detergents, samples for the structure and dynamics studies are fully functional.

PLN Structure and Dynamics in DPC Micelles

In the SR, PLN is thought to exist as an inactive pentamer (storage form), which depolymerizes into functional monomers prior to interaction with SERCA (Figure 2) (26, 27). Accordingly, we have focused our attention on a fully functional monomer of PLN (AFA-PLN) obtained by mutating the three cysteine residues (Cys-36, 41, 46) of the transmembrane domain to Ala, Phe, and Ala, respectively. Using NOE restraints in structural calculations, we have determined that PLN adopts an *L-shaped* conformation in DPC micelles comprised of three distinct structural domains: cytoplasmic domain Ia (helical from residues 2-16), loop (β-turn from residues 17-21), and transmembrane domain (helical from residues 22-52) (Figure 4A and B) (28). From simulated annealing calculations, we obtained a structural ensemble with very good convergence for each single structural domain (Figure 4A), but with the PLN conformers displaying an interhelical angle (angle between cytoplasmic and transmembrane domains) of 80 ± 22° (Figure 4B). Indeed, the limited number of NOEs detected in the loop did not allow us to constrain the two helical domains and obtain an ensemble of structures with a low RMSD over the entire protein backbone. Because of the lack of convergence on the orientation of the helical domain Ia with respect to the surface of the micelle, we used Mn^{2+} and 5' and 16'-doxyl stearic acids as paramagnetic probes of the topological arrangement of PLN in the micelle (28). Our results show that Mn^{2+} ions cause a

reduction in the resonance intensities of residues located at both termini, the loop, and Ser-16 and Thr-17, showing that both phosphorylation sites are solvent exposed. Both 5' and 16'-doxyl stearic acids caused a reduction in the resonance intensities of residues located in the micellar region. 16'-doxyl stearic acid affected residues in the core of the micelle (residues 35-45), as well as Leu-7, located in the middle of domain Ia, a residue that is likely buried in the hydrophobic region of the micelle. At higher 16'-doxyl stearic acid/PLN ratios, domain Ia is also considerably affected by the paramagnetic center, with Ala-11 displaying slightly reduced resonance intensity.

Our structural ensemble of the PLN monomer is in qualitative agreement with studies carried out in organic solvents (29, 30), where the authors found an overall *L-shaped* structure of PLN with the intervening loop in either a short flexible turn or a β-turn type III conformation. However, we found an orientation of the cytoplasmic helix, which is in better agreement with the amphipathicity of the PLN sequence (28). The most relevant PLN conformations in PDB 1N7L are those in which the hydrophobic side chains in domain Ia (Val-4, Leu-7, Ala-11, and Ala-15) are oriented toward the interior of the micelle and the hydrophilic residues point toward the bulk solvent, rendering the two phosphorylation sites exposed for interaction with their respective protein kinases.

While our NMR structural studies identified *three structural domains*, nuclear spin relaxation measurements and solvent accessibility experiments carried out on uniformly ^{15}N labeled PLN(31, 32) further subdivided the helical segment from 22-52 into two domains: domain Ib (residues 22-31), which is more dynamic and constituted by hydrophilic residues, and domain II (32-52), more hydrophobic and motionally

restricted. Therefore, we can divide PLN into *four dynamic regions*, which are characterized by different order parameters (see Figure 4A and B). Carr-Purcell-Meiboom-Gill (CPMG) based relaxation measurements also indicated the presence of slow dynamics (μs-ms motion) in domain Ia, the loop, and domain Ib (31, 32). While the slow dynamics of domain Ia and the loop were predicted based on the H/D exchange factors, the flexibility of residues within domain Ib was unexpected. The plasticity of this region supports biological evidence for the great importance of domain Ib and the molding necessary to *fit* into the binding groove of SERCA (20, 33).

Recently, a structure of wild-type PLN (wt-PLN) was reported showing the pentamer to be in a *bellflower* arrangement, where the cytoplasmic domain helix of each monomer makes an ~20° angle with respect to the bilayer normal (34), substantially different from the topology we reported for monomeric AFA-PLN (28). Our recent studies using solution NMR and EPR in DPC detergent micelles also point toward a dominating *L-shaped* conformation (or *pinwheel* model) for the pentamer (35). We also looked at the topology of pentameric wt-PLN in lipid bilayers using solid-state NMR, finding unambiguous evidence in support of the *pinwheel* model (see below).

SLN Structure and Dynamics in Micelles

The initial structure of SLN was determined in sodium dodecyl sulfate (SDS) micelles using a synthetic polypeptide (36). Under these experimental conditions, SLN has an α-helical conformation from residue F9 to R27 with RMSDs of 0.65 and 1.66 Å for backbone and side chain atoms, respectively. Both the N-terminus (M1 to L8) and the C-terminus (S28 to Y31) were found to be unstructured. Subsequently, we

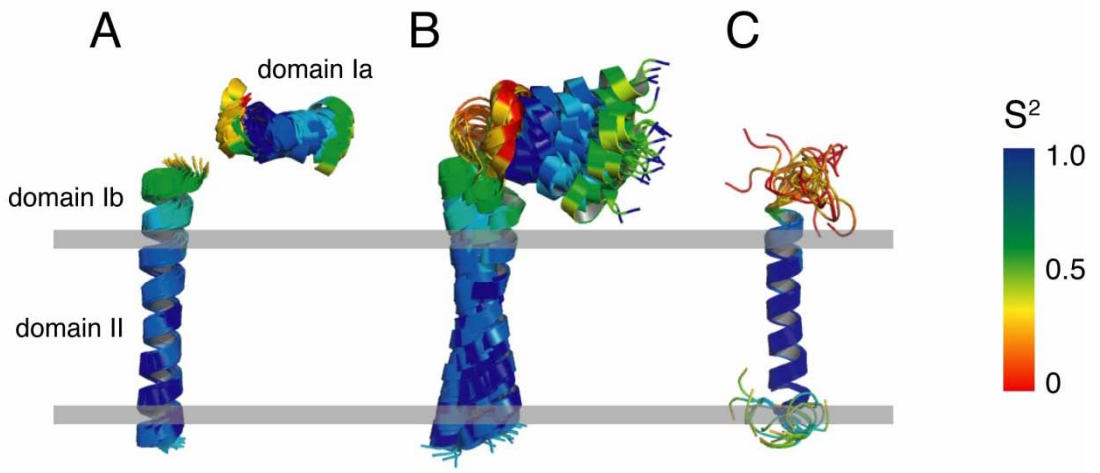


Figure 4. Structural ensembles of PLN and SLN from solution NMR. A) Separate overlay of PLN residues 2-16 and 22-52 from PDB *1N7L* (28). B) Entire backbone overlay of PLN from PDB *1N7L* (28). C) Backbone overlay of SLN as reported by Buffy *et al.* (24). The color-coding on the structures corresponds to the order parameters (S^2) as previously determined (31, 32).

expressed uniformly ^{15}N labeled recombinant SLN in *Escherichia coli*, which enabled the use of higher resolution [^{1}H, ^{15}N] NMR experiments in order to reduce several ambiguities in resonance overlap. With recombinant SLN, we determined the structure in DPC micelles, conditions that ensured the activity of SERCA, and found it to remain a single transmembrane helix with approximately five unstructured residues at either terminus (24). The superposition of the C_α, N_H and C' backbone atoms from residue R6 to R27 gave an RMSD of 0.4 ± 0.2 Å with an RMSD of 1.7 ± 0.3 Å for side chain atoms.

In addition to structure determination, we also measured spin relaxation rates, and found that the backbone dynamics, similar to PLN, is more complex than the structure. Relaxation measurements reveal four dynamic domains: a short unstructured N terminus (residues 1–6), a short dynamic helix (residues 7–14), a more rigid helix (residues 15–26), and an unstructured C terminus (residues 27–31) (24). H/D exchange factors also support the existence of four dynamic domains (24). The similarity of SLN structure and dynamics with those of PLN domains Ib and II shows that sequence conservation is reflected in the conservation of both structure and dynamics (Figure 4C).

PLN and SLN Topologies and Dynamics in Lipid Bilayers

Solid-state NMR in lipid bilayers has emerged as a complement to solution NMR studies in detergent micelles for elucidating structure, dynamics, and interactions between membrane proteins (37, 38). For our first solid-state NMR studies, we synthesized AFA-PLN ^{15}N labeled at Ala-11, Ala-15, and Ala-36, and measured the ^{15}N

chemical shift anisotropy in 4/1 DOPC/DOPE mechanically oriented lipid bilayers on glass plate supports (39). Using these solid-state NMR measurements in concert with rigid body molecular mechanics, we found that the domains Ib and II are oriented approximately perpendicular to the plane of the bilayer with the interhelical (i.e. interdomain) angle ranging between 60 and 100°, ruling out the possibility of a continuous α-helix and also suggesting that the cytoplasmic domain of PLN interacts with the membrane surface.

Similar solid-state NMR measurements on SLN oriented in mechanically aligned DOPC/DOPE bilayers revealed the approximate parallel orientation of the SLN helix with respect to the membrane bilayer normal (36). Since the limited number of labeled sites did not allow us to give quantitative topological angles for SLN and PLN within the bilayer, we then proceeded to use 2D [^{1}H, ^{15}N]-PISEMA (polarization inversion spin exchange at the magic angle) experiments (40). This separated-local-field experiment correlates the ^{15}N chemical shift anisotropy (CSA) with the ^{1}H-^{15}N dipolar coupling (DC). Since the values of both CSA and DC depend on the orientation of the peptide plane with respect to the direction of the magnetic field, the assignment of the amide resonances allows for the determination of the structure and topology in aligned lipid bilayers.

We determined that AFA-PLN has an overall *L-shaped* conformation in mixed 4/1 DOPC/DOPE lipid bilayers, where the helix comprising domains Ib and II makes a tilt angle of ~21° with respect to the bilayer normal (41). As expected from the 1D solid-state NMR studies on synthetically ^{15}N labeled AFA-PLN (36), PISEMA NMR

spectroscopy clearly shows that domain Ia interacts with the membrane surface, making an angle of ~93° with respect to the bilayer normal (41). A current model of the PLN monomer is reported in Figure 5.

In addition to the structural and topological information obtained from oriented alignments, tilting the aligned samples to different angles with respect to the direction of the static field makes it possible to investigate the rotational dynamics of the protein within the bilayer (42, 43, 44). Tilting the AFA-PLN sample by 90° revealed that the domains Ib and II undergo fast long-axial rotational diffusion about the bilayer normal with the cytoplasmic domain undergoing this same motion and other complex dynamics, scaling both the values of CSA and DC (41). The dynamics detected in both our solution and solid-state NMR experiments may explain variability within the literature regarding the topology of the cytoplasmic domain of PLN. For example, a magic-angle-spinning (MAS) solid-state NMR study carried out by Baldus and co-workers found that while cross-polarization (CP) based pulse sequences were adequate to detect the transmembrane domain of AFA-PLN, showing the existence of a well-defined helix, the cytoplasmic domain residues were too dynamic to be detected (45). Instead, J-coupling coherence transfers, similar to those in solution NMR experiments, were used to detect the dynamic cytoplasmic domain, which resulted in the conclusion that the cytoplasmic domain was completely unstructured. While this study represents advancement in MAS methodology, the structure most likely represents a minor conformational state and is inconsistent with a wealth of data, including those from our laboratory, which consistently show a predominant helical cytoplasmic domain with an overall *L-shaped* monomeric structure in lipid bilayers and detergent micelles.

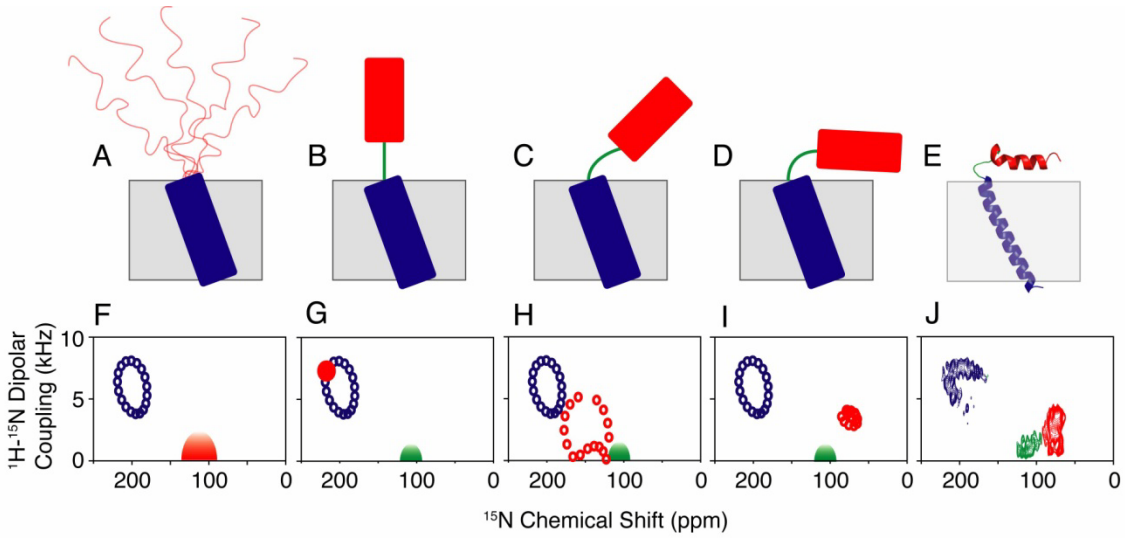


Figure 5: Structural models (panels A-D) proposed for PLN in lipid bilayers with PISEMA simulations (panels F-I). The ensemble of conformers reported by our laboratory (PDB 1N7L) is shown in panel E (28). The colors are coded with the regions of the protein: red for cytoplasmic, blue for transmembrane, and green for loop regions. The L-shaped topology (panels D and E) agrees best with the experimental PISEMA (panel J). Taken from Traaseth *et al.* (41).

A close inspection of PISEMA spectra from selectively labeled samples reveals the presence of two peak populations that exemplify two slightly different topologies for the AFA-PLN domains Ib and II (41). The two topologies have the same tilt angle (θ) for domain Ib and II with respect to the membrane normal, but slightly different rotational angles around the helix axis (ρ). Multiple populations of PLN have also been observed by the Lorigan and Middleton groups using MAS NMR experiments in lipid vesicles (46, 47). The detection of multiple conformers underscores the plasticity of PLN, and might be an important recognition mechanism for SERCA, protein kinase A, and protein phosphatase 1, previously shown necessary for physiological processes (33).

As with AFA-PLN, the PISEMA spectra of SLN obtained on uniformly ^{15}N labeled and selectively [^{15}N-Leu], [^{15}N-Ile], and [^{15}N-Val] samples also revealed the existence of two distinct topologies (48). Both the major and the minor populations of the resonances corresponding to domains Ib and II are oriented ~23° with respect to the lipid bilayer normal, but vary in the rotation angle about the helical axis by ~5° (in remarkable agreement with AFA-PLN). The primary sequence homology between SLN and PLN results in nearly identical structural and dynamic properties of these two regulatory proteins.

Pentameric wt-PLN Topology in Lipid Bilayers

More recently, our group has embarked on the validation of the pentameric structure of wt-PLN in lipid bilayers and detergent micelles (35). While monomeric PLN has previously been shown to bind and inhibit SERCA, a recent hypothesis

proposed that the pentamer could also bind and inhibit SERCA (49). While there is broad consensus regarding the secondary structure of penameric wt-PLN, there is disagreement in the literature over the orientation of the cytoplasmic helix. In particular, there are four proposed models for pentameric wt-PLN (Figure 6). The first model (*extended helix/sheet*) shows wt-PLN to be comprised of two α-helices connected by an anti-parallel β-sheet (residues 22-32), where the cytoplasmic domain is oriented 50-60° relative to the bilayer normal (50). The second model depicts wt-PLN as a continuous α-helix with a tilt angle of 28 ± 6° with respect to the bilayer normal (51, 52). The third model (*pinwheel*) shows that the most stable pentamer has a *pinwheel* geometry in which the cytoplasmic domain helices are oriented ~90° with respect to the membrane bilayer normal (53). The fourth and most recent model (*bellflower*) shows the structure of the pentamer to be in a *bellflower* assembly with the cytoplasmic domain helices oriented ~20° with respect to the bilayer normal (34).

To study the topology of pentameric wt-PLN, we reconstituted the protein in mechanically aligned 4/1 DOPC/DOPE lipid bilayers and analyzed the protein's architecture using [^{1}H, ^{15}N] PISEMA spectroscopy. As with the AFA-PLN monomer (41), we found that the wt-PLN PISEMA spectrum is composed of three different populations of resonances (see 1D spectrum in Figure 7A) corresponding to: a domains Ib and II (with the resonances located between 170-220 ppm), an in-plane cytoplasmic domain (with resonances located between 50-100 ppm), and a more flexible region (loop and termini) with resonances clustered around ~110 ppm (isotropic portion of the spectrum). Since our 1D spectrum shows three distinct regions, indicating three unique

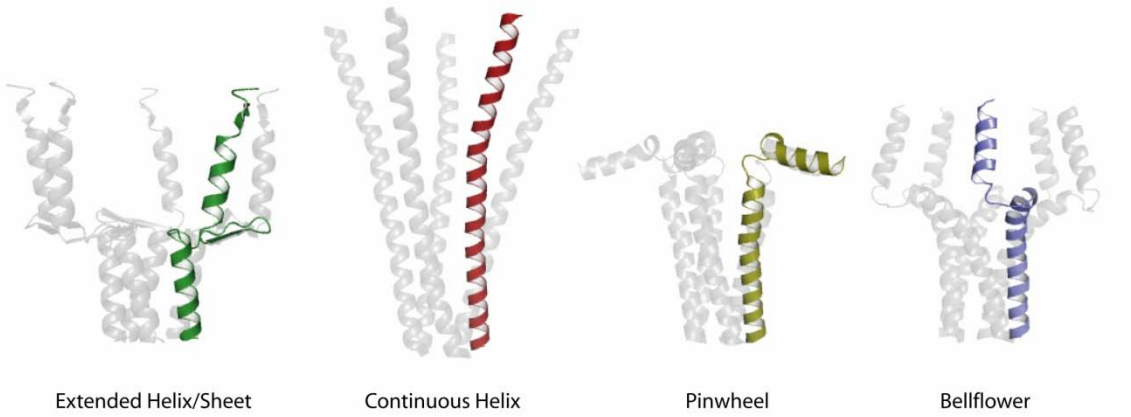


Figure 6: Structural models of wt-PLN. The *pinwheel* (*1XNU*) and *bellflower* (*1ZLL*) pentamer models were taken directly from PDB coordinates. Taken from Traaseth *et al.* (35).

wt-PLN domain alignments with respect to the membrane bilayer normal, this eliminates the possibility of the *continuous helix* model.

To distinguish between the other models depicted in Figure 6, we performed PISEMA experiments using selectively [^{15}N-Ala], [^{15}N-Thr], [^{15}N-Leu], [^{15}N-Ile], [^{15}N-Cys], and [^{15}N-Asn] labeled wt-PLN samples. Our experimental PISEMA spectra, reported as an overlay in Figure 7, show a remarkable similarity with the AFA-PLN monomer. In fact, the simulations for the cytoplasmic domain resonances (Leu-7, Thr-8, Ala-11, Ile-12, Ala-15, Thr-17, Ile-18) correspond to a helix with a tilt angle of ~90° with respect to the bilayer normal (Figure 7E) (35).

Figure 8 shows PISEMA spectra simulated from the *pinwheel* and *bellflower* PDB coordinates for those selectively labeled sites shown in Figure 7. If the pentamer topology corresponded to the *pinwheel* model, the cytoplasmic domain residues would resonate in the upfield region of the spectrum (50-100 ppm) (Figure 8A). On the other hand, if the architecture of wt-PLN were consistent with the *bellflower* model, the cytoplasmic domain resonances in the PISEMA pattern would occupy the downfield portion of the spectrum (170-220 ppm), as represented in Figure 8B. Comparing the experimental cytoplasmic domain spectra, it is clear that in lipid bilayers the cytoplasmic domain is oriented perpendicular with the bilayer normal forming an overall *pinwheel* geometry.

Structural fitting with an ideal helix in Figure 7D revealed that the helix corresponding to domain Ib and II of pentameric wt-PLN has a tilt angle (θ) of ~15° with respect to the bilayer normal (35). Monomeric AFA-PLN has a tilt angle of ~21°

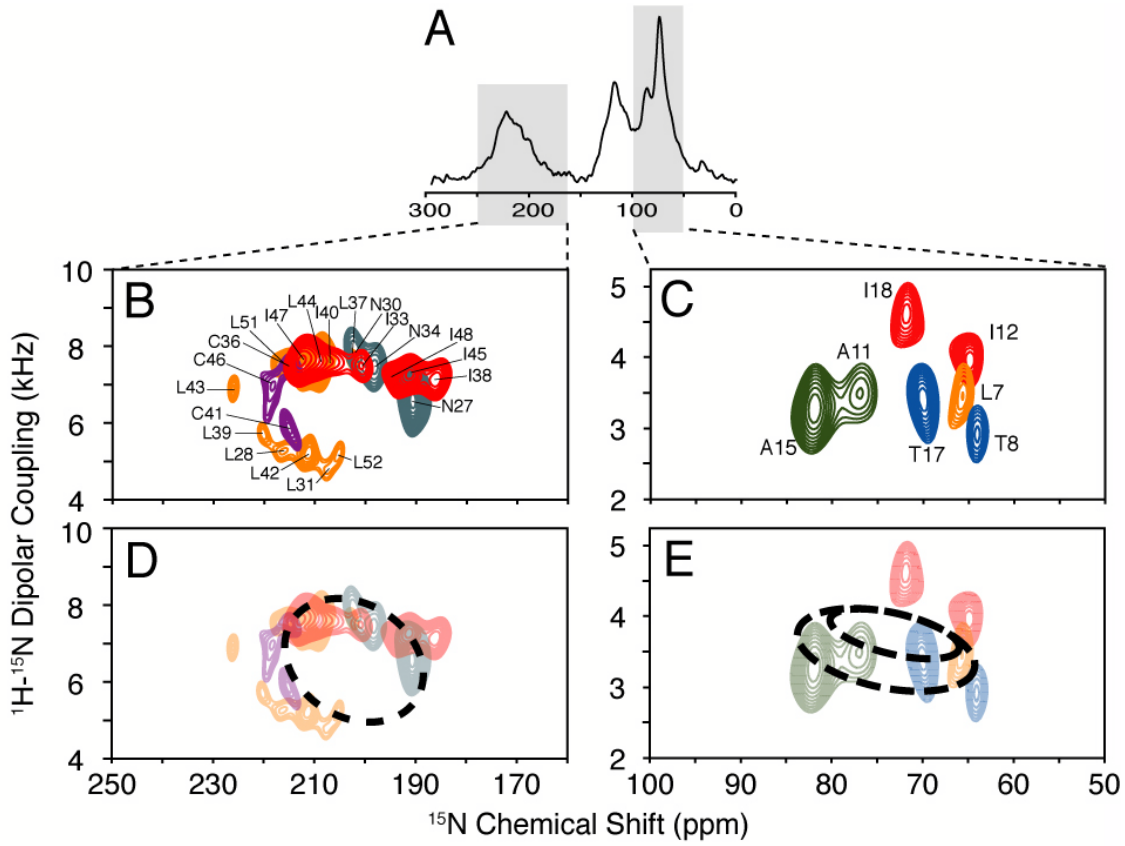


Figure 7: PISEMA spectra of PLN pentamer in lipid bilayers. (A) 1D cross-polarization spectrum of [U-^{15}N] wt-PLN in DOPC/DOPE oriented lipid bilayers. (B, C) Selectively labeled wt-PLN PISEMA spectra for the transmembrane and cytoplasmic helices, respectively. The residues are color coded with the PISEMA spectra: [^{15}N-Ala] green, [^{15}N-Cys] purple, [^{15}N-Leu] orange, [^{15}N-Ile] red, [^{15}N-Asn] gray, [^{15}N-Thr] blue. (D, E) Simulated PISA wheels for both transmembrane (θ=15°) and cytoplasmic (θ=92°) domains. Taken from Traaseth *et al.* (35).

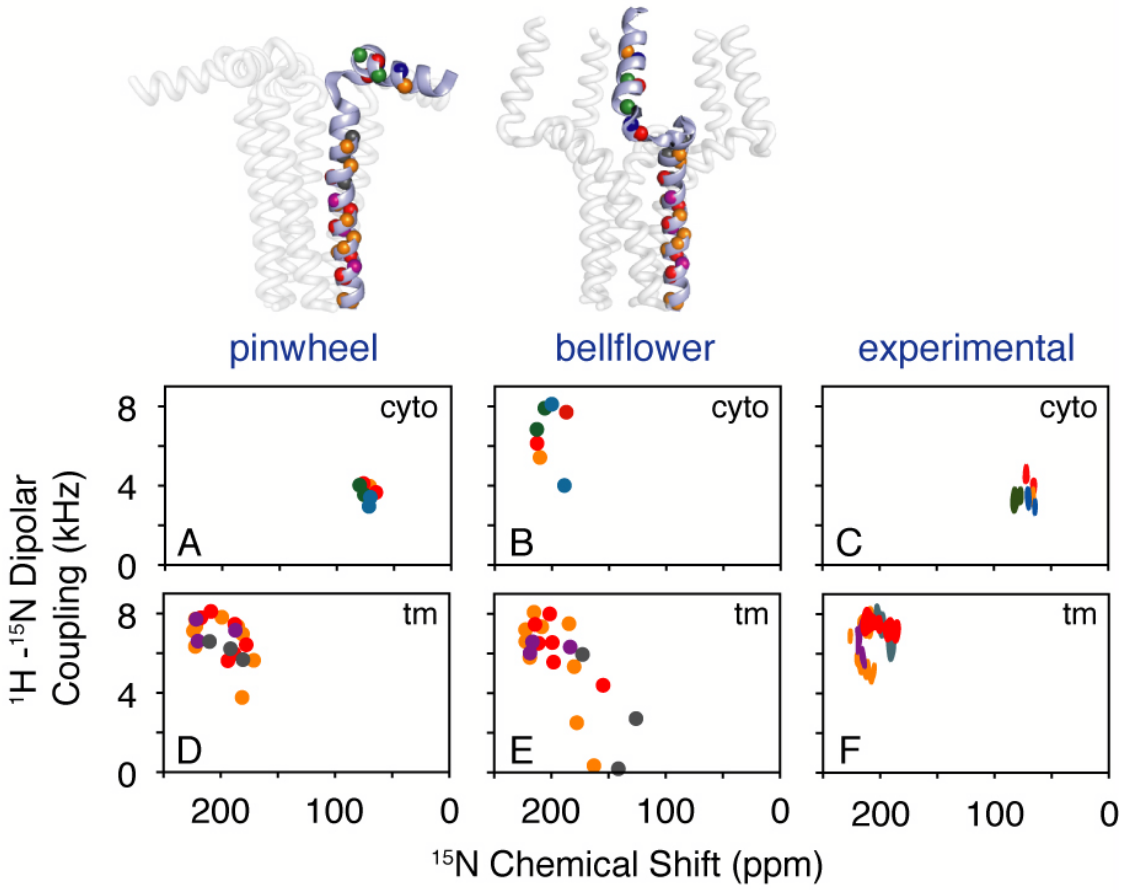


Figure 8: Simulated PISEMA spectra for the pinwheel and bellflower models. Unlike the *pinwheel* model, the *bellflower* model shows no high-field resonances. Experimental PISEMA spectra show the remarkable agreement with the *pinwheel* model (C, F). Taken from Traaseth *et al.* (35).

(41), which requires pentamer formation to tilt by ~6° to accommodate the leucine/isoleucine zipper holding the pentamer together (54, 55).

Allosteric Activation Model

The functionality of SERCA under NMR conditions and the quality of AFA-PLN spectra upon addition of SERCA enabled the *unprecedented* atomic mapping of the interactions between these two integral membrane proteins in detergent micelles (56). In its free form, AFA-PLN exists in a dynamic equilibrium between two conformations, T and R states, where the T state or *L-shaped* conformation is thermodynamically stable, and the R state or extended form is identified with a more dynamic cytoplasmic domain (Figure 9) (56, 57). These two states are readily detected using EPR spectroscopy in both micelles and lipid bilayers (56, 57), but due to the timescale of the exchange NMR can only imply the existence of these forms from relaxation dispersion measurements (i.e. conformational interconversion). However, upon addition of SERCA to AFA-PLN, chemical shift perturbation analyses reveal the appearance of a second population of peaks within domain Ia, the loop, and domain Ib, indicating a conformational switch of AFA-PLN from the T to the R state, a process exemplifying an allosteric activation mechanism (56).

Resonances from the hydrophobic portion of the transmembrane region (domain II) also show chemical exchange to the R state (25). A difference plot of $^{1}H_{N}$ chemical shift before and after addition of SERCA for residues 32-52 shows a symmetric bimodal behavior where the C-terminal part of domain II shifted upfield and the residues near the N-terminal part downfield (25). Since upfield and downfield shifts

have been correlated to the strength of hydrogen bonds (58), one possible explanation of the data is that the C-terminal end of the transmembrane domain (residues 46-52) unwinds upon binding SERCA. This hypothesis was first proposed by MacLennan and co-workers, who indicated an overall change in the secondary structure of the transmembrane domain, with residues 49-52 unwinding upon interaction with SERCA, a process that might facilitate binding (20).

These results are echoed in the binding of SLN to SERCA (24). Overall, SLN behaves like the transmembrane domain of AFA-PLN, with each dynamic domain mimicking the behavior of the corresponding domain in AFA-PLN. Upon addition of SERCA, the transmembrane domain is in fast exchange between two free forms (T and R states). As previously indicated, spin relaxation measurements dissected the transmembrane domain of SLN into two regions we named domain Ib and domain II in analogy with AFA-PLN. The chemical shift changes of these two regions follow the bimodal behavior of the transmembrane domain of AFA-PLN, indicative of a similar mechanism involving an unwinding of the C-terminal residues and a stabilization of the residues in the N-terminal portion of the protein as a result of the interactions with SERCA (24). This supports the hypothesis that both SLN and PLN transmembrane domains bind SERCA in the same site and with an identical mechanism.

Effects of PLN Phosphorylation on the Allosteric Mechanism

The inhibition of SERCA by PLN can be reversed by phosphorylation at Ser-16 by cAMP-dependent protein kinase A (1). We solved the structure of Ser-16

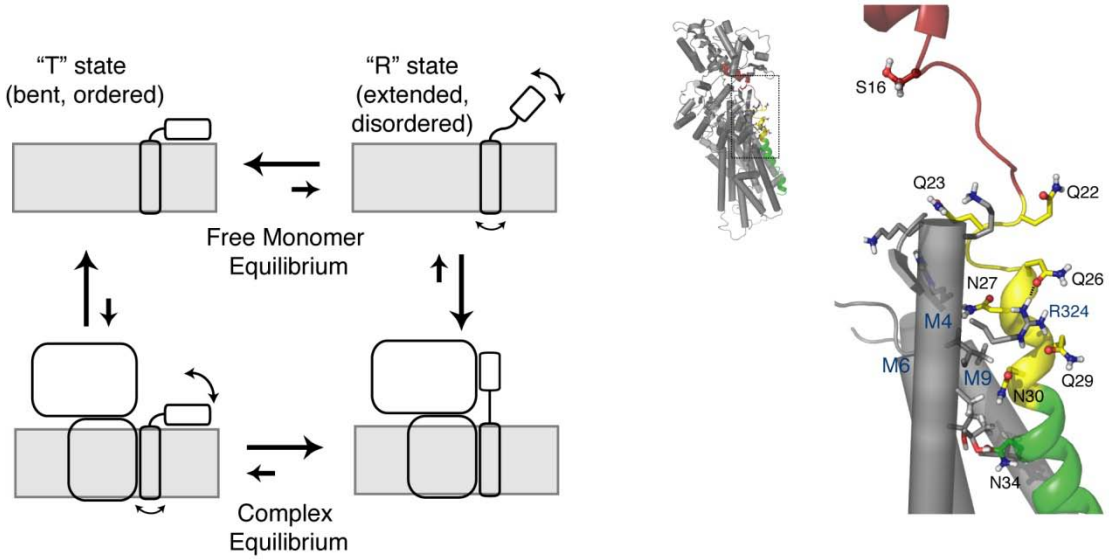


Figure 9. Allosteric model of SERCA inhibition and PLN/SERCA structural model. A) Allosteric model of PLN interaction with SERCA and the effect of phosphorylation at Ser-16 (indicated in red). PLN monomer interconverts between the L-shaped form (T state) and the less stable (more dynamically disordered) extended form (R state). B) PLN/SERCA model developed by Toyoshima *et al.* (20) highlighting the long range allosteric control phosphorylation at Ser-16 has on domain Ib. Figure adapted from Zamoon *et al.* (56) and Traaseth *et al*. (25).

phosphorylated AFA-PLN (pS16-AFA-PLN) and found that residues 14-16, previously helical, became unwound upon phosphorylation, revealing an order-to-disorder transition (32). In addition, we found that there are pronounced changes in pS16-AFA-PLN backbone dynamics on both the ps-ns and μs-ms timescales (32). Although small, some of the changes are propagated throughout the entire protein backbone, demonstrating that while the structural transitions following phosphorylation are localized, the changes in backbone dynamics are irradiated throughout the protein.

How can this order-to-disorder transition help in understanding the interaction with SERCA? To answer this question, we proceeded with the analysis of chemical shift perturbation of pS16-AFA-PLN induced by SERCA (25). We found that the conformational equilibrium between the T and R states upon addition of SERCA is influenced by the single phosphorylation at Ser-16; specifically, phosphorylation shifts the equilibrium toward the R state in a cooperative manner (25). Another considerable difference upon phosphorylation includes a change in both the surface and the dynamics of domain Ib. In particular, a remarkable change is observed for the side-chain binding behavior. In constrast with unphosphorylated PLN, the Gln-26 resonance in the phosphorylated protein is unperturbed by SERCA with other smaller changes seen for side chain residues Asn-27, 30, and 34. A possible mechanism to explain these results is the rotation or rearrangement of domain Ib upon phosphorylation that disrupts crucial intermolecular hydrogen bonds, resulting in relief of inhibition. Based on the molecular model by MacLennan and co-workers (20), we proposed that a crucial hydrogen bond formed between Arg-324 and Gln-26 may be broken between AFA-PLN and SERCA after phosphorylation at Ser-16 (25). These findings are in agreement with mutagenesis

studies, showing that Q26A is a loss-of-function mutant (59). In addition, cross-linking studies show that an N27C mutation in pS16-AFA-PLN is no longer able to cross-link with SERCA (20). While there are changes in the binding interface for domain Ib, the overall binding of the transmembrane domain to SERCA is not affected by phosphorylation. The overall dissociation constants (K_d) for domain II in both pS16-AFA-PLN and AFA-PLN are ~60 μM (25). This demonstrates that the major changes are in domain Ia, loop, and domain Ib, with domain II only marginally affected, supporting the hypothesis that phosphorylation at Ser-16 does not dissociate PLN from SERCA completely. We proposed domain Ib as a bridgehead region, which transmits the dynamics induced by phosphorylation at Ser-16 from the cytoplasmic helix to domain Ib, thereby regulating the intramembrane protein-protein interaction (25). A schematic of the allosteric model for phosphorylation is reported in Figure 9.

Perspective

What can we learn from the analysis of the structure and dynamics of PLN? More importantly, how are structure and dynamics of PLN correlated to SERCA's function, and can we control the extent of inhibition of SERCA by manipulating PLN structural dynamics? These are questions we have begun to address concerning PLN and look forward to answering in the upcoming years.

Given the plethora of biochemical and molecular biology data currently available, it is a very exciting moment for the structural biologists involved in research on PLN and SLN and their interactions with SERCA. While our studies to date have focused on detecting the effects of PLN induced by SERCA, our future challenge

involves detecting SERCA changes from PLN within the entire enzymatic cycle. Another important challenge is to study the SERCA/PLN complex in the complicated network of interactions involving protein kinase A and protein phosphatase 1. While solution NMR will help identify some important pieces of this complex puzzle, solid-state NMR will be the method of choice for the elucidation of the structural dynamics and interactions in these large complexes.

References

(1) Wegener, A. D., Simmerman, H. K., Lindemann, J. P., and Jones, L. R. (1989) *J. Biol. Chem. 264*, 11468-11474.

(2) Mundina-Weilenmann, C., Vittone, L., Ortale, M., de Cingolani, G. C., and Mattiazzi, A. (1996) *J. Biol. Chem. 271*, 33561-33567.

(3) Chu, G., Lester, J. W., Young, K. B., Luo, W., Zhai, J., and Kranias, E. G. (2000) *J. Biol. Chem. 275*, 38938-38943.

(4) Wawrzynow, A., Theibert, J. L., Murphy, C., Jona, I., Martonosi, A., and Collins, J. H. (1992) *Arch. Biochem. Biophys. 298*, 620-623.

(5) Odermatt, A., Taschner, P. E., Scherer, S. W., Beatty, B., Khanna, V. K., Cornblath, D. R., Chaudhry, V., Yee, W. C., Schrank, B., Karpati, G., Breuning, M. H., Knoers, N., and MacLennan, D. H. (1997) *Genomics. 45*, 541-53.

(6) Gayan-Ramirez, G., Vanzeir, L., Wuytack, F., and Decramer, M. (2000) *J. Physiol. 524 Pt 2*, 387-397.

(7) Minamisawa, S., Wang, Y., Chen, J., Ishikawa, Y., Chien, K. R., and Matsuoka, R. (2003) *J. Biol. Chem. 278*, 9570-9575.

(8) Babu, G. J., Bhupathy, P., Carnes, C. A., Billman, G. E., and Periasamy, M. (2007) *J. Mol. Cell. Cardiol. 43*, 215-222.

(9) MacLennan, D. H., Yip, C. C.,Iles, and Seeman, P. (1972) *Cold Spring Harbor Symp. Quant. Biol. 37*, 469-78.

(10) Odermatt, A., Becker, S., Khanna, V. K., Kurzydlowski, K., Leisner, E., Pette, D., and MacLennan, D. H. (1998 *J. Biol. Chem. 273*, 12360-12369.

(11) Uemura, N., Ohkusa, T., Hamano, K., Nakagome, M., Hori, H., Shimizu, M., Matsuzaki, M., Mochizuki, S., Minamisawa, S., and Ishikawa, Y. (2004) *Eur. J. Clin. Invest. 34*, 723-730.

(12) Gramolini, A. O., Trivieri, M. G., Oudit, G. Y., Kislinger, T., Li, W., Patel, M. M., Emili, A., Kranias, E. G., Backx, P. H., and Maclennan, D. H. (2006) *Proc. Natl. Acad. Sci. U. S. A. 103*, 2446-2451.

(13) Toyoshima, C., Nakasako, M., Nomura, H., and Ogawa, H. (2000) *Nature. 405*, 647-55.

(14) Toyoshima, C., and Nomura, H. (2002) *Nature. 418*, 605-11.

(15) Toyoshima, C., and Mizutani, T. (2004) *Nature. 430*, 529-535.

(16) Toyoshima, C., Nomura, H., and Tsuda, T. (2004) *Nature. 432*, 361-368.

(17) Olesen, C., Sorensen, T. L., Nielsen, R. C., Moller, J. V., and Nissen, P. (2004) *Science. 306*, 2251-2255.

(18) Sorensen, T. L., Moller, J. V., and Nissen, P. (2004) *Science. 304*, 1672-1675.

(19) Young, H. S., Jones, L. R., and Stokes, D. L. (2001) *Biophys J. 81*, 884-94.

(20) Toyoshima, C., Asahi, M., Sugita, Y., Khanna, R., Tsuda, T., and MacLennan, D. H. (2003) *Proc Natl Acad Sci U S A. 100*, 467-72.

(21) Asahi, M., Sugita, Y., Kurzydlowski, K., De Leon, S., Tada, M., Toyoshima, C., and MacLennan, D. H. (2003) *Proc Natl Acad Sci U S A. 100*, 5040-5.

(22) Hutter, M. C., Krebs, J., Meiler, J., Griesinger, C., Carafoli, E., and Helms, V. (2002) *Chembiochem. 3*, 1200-8.

(23) Buck, B., Zamoon, J., Kirby, T. L., DeSilva, T. M., Karim, C., Thomas, D., and Veglia, G. (2003) *Protein Expr Purif. 30*, 253-61.

(24) Buffy, J. J., Buck-Koehntop, B. A., Porcelli, F., Traaseth, N. J., Thomas, D. D., and Veglia, G. (2006) *J. Mol. Biol. 358*, 420-429.

(25) Traaseth, N. J., Thomas, D. D., and Veglia, G. (2006) *J. Mol. Biol. 358*, 1041-1050.

(26) Kimura, Y., Kurzydlowski, K., Tada, M., and MacLennan, D. H. (1997) *J Biol Chem. 272*, 15061-4.

(27) Reddy, L. G., Jones, L. R., and Thomas, D. D. (1999) *Biochemistry. 38*, 3954-62.

(28) Zamoon, J., Mascioni, A., Thomas, D. D., and Veglia, G. (2003) *Biophys J. 85*, 2589-98.

(29) Pollesello, P., Annila, A., and Ovaska, M. (1999) *Biophys J. 76*, 1784-95.

(30) Lambeth, S., Schmid, H., Muenchbach, M., Vorherr, T., Krebs, J., Carafoli, E., and Griesinger, C. (2000) *Chimica Acta. 83*, 2141-2152.

(31) Metcalfe, E. E., Zamoon, J., Thomas, D. D., and Veglia, G. (2004) *Biophysical Journal. 87*, 1-10.

(32) Metcalfe, E. E., Traaseth, N. J., and Veglia, G. (2005) *Biochemistry. 44*, 4386-96.

(33) Schmidt, A. G., Zhai, J., Carr, A. N., Gerst, M. J., Lorenz, J. N., Pollesello, P., Annila, A., Hoit, B. D., and Kranias, E. G. (2002) *Cardiovasc. Res. 56*, 248-259.

(34) Oxenoid, K., and Chou, J. J. (2005) *Proc. Natl. Acad. Sci. U. S. A. 102*, 10870-10875.

(35) Traaseth, N. T., Verardi, R., Torgersen, K. D., Karim, C. B., Thomas, D. D., and Veglia, G. (2007) *PNAS. In Press*.

(36) Mascioni, A., Karim, C., Barany, G., Thomas, D. D., and Veglia, G. (2002) *Biochemistry. 41*, 475-82.

(37) Opella, S. J. (1997) *Nat Struct Biol. 4 Suppl*, 845-8.

(38) Baldus, M. (2006) *Curr. Opin. Struct. Biol. 16*, 618-623.

(39) Mascioni, A., Karim, C., Zamoon, J., Thomas, D. D., and Veglia, G. (2002) *J Am Chem Soc. 124*, 9392-3.

(40) Wu, C. H., Ramamoorthy, A., and Opella, S. J. (1994) *J Mag Res. 109*, 270-272.

(41) Traaseth, N. J., Buffy, J. J., Zamoon, J., and Veglia, G. (2006) *Biochemistry. 45*, 13827-13834.

(42) Grage, S. L., Wang, J., Cross, T. A., and Ulrich, A. S. (2002) *Biophys. J. 83*, 3336-3350.

(43) Aisenbrey, C., and Bechinger, B. (2004) *J. Am. Chem. Soc. 126*, 16676-16683.

(44) Park, S. H., Mrse, A. A., Nevzorov, A. A., De Angelis, A. A., and Opella, S. J. (2006) *J. Magn. Reson. 178*, 162-165.

(45) Andronesi, O. C., Becker, S., Seidel, K., Heise, H., Young, H. S., and Baldus, M. (2005) *J. Am. Chem. Soc. 127*, 12965-12974.

(46) Hughes, E., Clayton, J. C., and Middleton, D. A. (2005) *Biochemistry. 44*, 4055-4066.

(47) Karp, E. S., Tiburu, E. K., Abu-Baker, S., and Lorigan, G. A. (2006) *Biochim. Biophys. Acta. 1758*, 772-780.

(48) Buffy, J. J., Traaseth, N. J., Mascioni, A., Gor'kov, P. L., Chekmenev, E. Y., Brey, W. W., and Veglia, G. (2006) *Biochemistry. 45*, 10939-10946.

(49) Stokes, D. L., Pomfret, A. J., Rice, W. J., Glaves, J. P., and Young, H. S. (2006) *Biophys. J. 90*, 4213-4223.

(50) Tatulian, S. A., Jones, L. R., Reddy, L. G., Stokes, D. L., and Tamm, L. K. (1995) *Biochemistry. 34*, 4448-56.

(51) Arkin, I. T., Rothman, M., Ludlam, C. F., Aimoto, S., Engelman, D. M., Rothschild, K. J., and Smith, S. O. (1995) *J Mol Biol. 248*, 824-34.

(52) Smith, S. O., Kawakami, T., Liu, W., Ziliox, M., and Aimoto, S. (2001) *J Mol Biol. 313*, 1139-48.

(53) Robia, S. L., Flohr, N. C., and Thomas, D. D. (2005) *Biochemistry. 44*, 4302-4311.

(54) Simmerman, H. K., Kobayashi, Y. M., Autry, J. M., and Jones, L. R. (1996) *J Biol Chem. 271*, 5941-6.

(55) Karim, C. B., Stamm, J. D., Karim, J., Jones, L. R., and Thomas, D. D. (1998) *Biochemistry. 37*, 12074-81.

(56) Zamoon, J., Nitu, F., Karim, C., Thomas, D. D., and Veglia, G. (2005) *Proc Natl Acad Sci U S A. 102*, 4747-52. Epub 2005 Mar 21.

(57) Karim, C. B., Kirby, T. L., Zhang, Z., Nesmelov, Y., and Thomas, D. D. (2004) *Proc Natl Acad Sci U S A. 101*, 14437-42. Epub 2004 Sep 24.

(58) Wagner, G., Pardi, A., and Wuthrich, K. (1983) *J Am Chem Soc. 105*, 5948-9.

(59) Kimura, Y., Asahi, M., Kurzydlowski, K., Tada, M., and MacLennan, D. H. (1998) *J Biol Chem. 273*, 14238-41.

(60) Inesi, G., Lewis, D., Ma, H., Prasad, A., and Toyoshima, C. (2006) *Biochemistry. 45,* 13769-13778.