Technische Universität München

Fakultät für Chemie

Biomolekulare NMR Spektroskopie

Insights in the structure and function of the N-terminal regions of *Mycobacterium tuberculosis* protein kinase G by NMR and complementary methods

Matthias Wittwer

Für Helene

Für Helene

Abstract

Mycobacterium tuberculosis protein kinase G (PknG) is an 82 kDa multidomain eukaryotic - like serine / threonine kinase that mediates the survival of pathogenic mycobacteria in host macrophages. It has been proven that the N-terminal NORS region of PknG plays a role in PknG regulation by (auto)phosphorylation, whereas the following rubredoxin-like metal bound motif interacts with the catalytic domain and is involved in redox regulation. In addition, it was shown that deletions or mutations within both domains significantly reduce the cellular survival of mycobacteria. The objective of this study was, to elucidate the structure and the dynamic behavior of the N-terminal regions by using a combination of NMR spectroscopy, *in vitro* kinase assays and MD simulation data, to gain new insights into their regulatory role for PknG kinase activity. It was found that the NORS motif is highly flexible and intrinsically unstructured, whereas the following RD in its metal-bound state is more rigid and mostly well structured. Moreover, based on *in vitro* kinase assay and MD simulation data, we assume that substrate access to the catalytic domain is regulated by the controlled folding and unfolding of the RD and that PknG is able to respond to environmental redox changes caused by ROS within host macrophages.

Zusammenfassung

Mycobacterium tuberculosis protein kinase G (PknG) ist eine 82 kDa große Eukaryonten ähnliche multidomainen Serin / Threoninkinase, die das Überleben von pathogenen Mykobakterien in Wirtsmakrophagen vermittelt. Die N-terminale NORS-Region von PknG spielt nachweislich eine Rolle bei der Regulierung von PknG durch (Auto-)Phosphorylierung, während das folgende Rubredoxin-ähnliche Metallbindungsmotiv mit der katalytischen Domäne interagiert und an der Regulierung des Redoxzustandes beteiligt ist. Darüber hinaus wurde gezeigt, dass Deletionen oder Mutationen innerhalb beider Domänen das zelluläre Überleben von Mykobakterien signifikant reduzieren. Ziel dieser Studie war es, die Struktur und das dynamische Verhalten der N-terminalen Regionen mit Hilfe einer Kombination aus NMR-Spektroskopie, *in-vitro* Kinaseassays und MD-Simulationsdaten aufzuklären, um neue Erkenntnisse über ihre regulatorische Rolle für die Kinaseaktivität von PknG zu gewinnen. Es wurde festgestellt, dass das NORS-Motiv hochflexibel und intrinsisch unstrukturiert ist, während die folgende RD im metallgebundenen Zustand rigider und meist gut strukturiert ist. Darüber hinaus gehen wir davon aus, dass der Zugang des Substrats zur katalytischen Domäne durch die kontrollierte Faltung und Entfaltung der RD reguliert wird und das PknG in der Lage ist, auf Redoxänderungen in seiner Umwelt zu reagieren, die durch ROS in Wirtsmakrophagen verursacht werden.

Table of Contents

List of Figures

List of Tables

Abbreviations

ADP	Adenosine 5'-diphosphate
ATP	Adenosine triphosphate
CHAPSO	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
DHPC	Dihexanoylphosphatidylcholine
DIODPC	1,2-Di-O-Dodecyl-sn-Glycero-3-Phosphocholine
DMPC	Dimyristoylphosphatidylcholine
DNA	Deoxyribonucleic acid
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
EDTA	Ethylenediaminetetraacetic acid
eSTK	Eukaryotic like serine / threonine kinase
GarA	Glycogen accumulation regulator A
HIV	Human Immunodeficiency Virus
HMQC	Hetero Multiple Quantum Correlation
HSQC	Hetero Single Quantum Coherence
kDa	Kilo Dalton
LB	Lysogeny Broth (Luria-Bertani broth)
MD	Molecular Dynamic
MDR-TB	Multi Drug Resistant - Tuberculosis
MEAD	Macroscopic Electrostatics with Atomic Detail
mM	milli molar
Mtb	*Mycobacterium tuberculosis*
MWCO	Molecular Weight Cut Off
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect Spectrocopy
NORS	No Regulatory Secondary Structure
PEP	Preservation of Equivalent Pathways
PknG	Protein Kinase G
ppm	parts per million
RD	Rubredoxin Domain
RDC	Residual Dipolar Coupling
ROS	Reactive Oxygen Species
S/N	Signal to Noise

TB	Tuberculosis
TCEP	Tris(2-carboxyethyl)phosphine
TMS	Tetramethylsilane
TOCSY	Total Correlation Spectroscopy
TPRD	Tetratricopeptide repeat domain
TSP	Trimethylsilylpropanic acid

Publications

Wittwer M., Dames S. A.: Expression and purification of the natively disordered and redox sensitive metal binding regions of *Mycobacterium tuberculosis* protein kinase G. *Protein Expr. Purif.* (2015) **111** 68-74, PMID: 25839674

Wittwer M., Dames S. A.: Chemical shift assignment of the natively disordered N-terminus and the rubredoxin domain in the folded, metal bound and unfolded, oxidized state of mycobacterial protein kinase G**.** *Biomol. NMR Assign.* (2016) **10** (2) 401-406, PMID: 27632081

Wittwer M., Luo Q., Kaila V. R., Dames S. A.: Oxidative Unfolding of the Rubredoxin Domain and the Natively Disordered N-terminal Region Regulate the Catalytic Activity of *Mycobacterium tuberculosis* Protein Kinase G. J. Biol. Chem. (2016) **291** (53) 27062-27072, PMID: 27810897

Chapter I

1. Biological Background

Tuberculosis (TB) is an infectious bacterial disease caused by *Mycobacterium tuberculosis* (Mtb) and ranks alongside HIV as a leading cause of mortality worldwide. In 2014 more than 9.6 million people are estimated to have fallen ill with TB. Globally, an estimated 3.3 % of new TB cases are infected with multi drug resistant strains (MDR-TB).The increasing prevalence of MDR-TB demands the need for the development of novel treatment strategies [1-3]. The fact, that the virulence of pathogenic mycobacteria is coupled to their capacity to survive for extended periods of time within human macrophages, makes mycobacterial protein kinase G (PknG) an attractive and promising drug target [4-6]. It was shown that PknG is essential for intracellular mycobacterial survival by blocking phagosome lysosome fusion, and that its inactivation results in mycobacterial cell death [7].In addition to its role in Mtb pathogenicity, it was further demonstrated that PknG is also involved in the regulation of the glutamate metabolism through the *in vivo* phosphorylation of the Mtb glycogen accumulation regulator (GarA) [8-12].

PknG is one of eleven eukaryotic like serine / threonine kinase (eSTK) encoded in the Mtb genome [13,14]. As a multidomain soluble protein, PknG consists of four regulatory domains and has a total molecular weight of 82 kDa. The conserved canonical kinase domain is flanked C-terminally by a tetratricopeptide repeat domain (TPRD), a motif involved in mediating protein-protein interactions, and N-terminally preceded by a presumably natively disordered NORS (no regulatory secondary structure) region and a rubredoxin domain (RD) [15]. The redox sensitive RD is characterized by two conserved CXXCG motifs which can tetrahydrally coordinate a divalent metal ion, based on previous *in vitro* studies, this can be either zinc, iron, or cadmium [15-18]. The metal ion which is coordinated naturally or *in vivo* is still unknown. Mutations of all key cysteines, to alanine or serine impairs the kinase activity and renders PknG insensitive to a regulation by redox changes [15,19]. Furthermore, it was shown that a deletion of the RD results in a higher activity of the catalytic domain against a 17mer peptide substrate derived from the endogenous protein GarA. The authors propose that the RD regulates the intrinsic PknG kinase activity by restricting the accessibility of the active site [16]. Contrary to most other Mtb eSTK's, PknG is classified as a non-RD-kinase, which is accompanied by the lack of a regulatory S/T autophosphorylation residue within the activation segment of the catalytic core [20-22]. However, it could be shown that the N-terminal NORS region is autophosphorylated *in vivo* on the threonine residue T32 and that the phosphorylation does not affect kinase activity but is required for survival of mycobacteria within host macrophages [19,23].

However, the published crystal structures provide no insight into the structure as well as the dynamic behavior of the intrinsically disordered N-terminal NORS region in absence or presence of phosphorylated residues. Moreover, the exact mechanism for the regulation of the kinase activity under oxidative stress conditions through the redox sensitive RD has not been elucidated yet.

For a better understanding of how the catalytic activity of PknG is regulated by its N-terminal regions a protocol for the design, expression and purification of the NORS region as well as for the redox sensitive RD was established [18]. Furthermore, the structure and dynamics of the NORS region as well as for the RD in different redox states were characterized by NMR spectroscopy [24]. Complementary*, in vitro* kinase assays and MD simulations were performed to investigate how the RD regulates the accessibility to catalytic site of PknG under oxidizing or reducing conditions [25].

2. NMR Spectroscopy of Biomolecules

Nuclear magnetic resonance (NMR) spectroscopy is a powerful non-destructive analytical tool for the structural determination of biological macromolecules. As a versatile technique NMR provides insights into motional and molecular information at atomic resolution. Its broad field of application allows the study of unique and essential dynamic and interaction profiles, containing pivotal information on protein folding and function. Molecular details obtained from NMR data assist target-directed drug discovery and design.

2.1 Basic principles

For a more comprehensive understanding of the physical principles of NMR, the reader is referred to the literature [26-31].

2.1.1 Nuclear Spin

Nuclear magnetic resonance (NMR), is a physical phenomenon based on the fact that nuclei of atoms posses an intrinsic angular momentum $\vec{I}$, known as spin. The nuclear spin is specified by the spin quantum number I which can be integer (0, 1, 2....) or half integer ($\frac{1}{2}, \frac{3}{2}$..). For NMR spectroscopy of biomolecules, the most important nuclei with spin $\frac{1}{2}$ are ^{1}H, ^{15}N, ^{13}C, ^{19}F, ^{31}P and ^{2}H, ^{14}N for nuclei with spin 1, respectively. Atomic nuclei with a zero-spin quantum number I, such as ^{12}C, are not detectable. The intrinsic angular momentum $\vec{I}$ for the z-direction can be described as follows (Eq. 1.1),

$$\vec{I_z} = \hbar m \qquad \text{Eq. 1.1}$$

where $\hbar$ is the reduced Planck constant ($\hbar$=h/2π) and m is the magnetic quantum number, whose value can vary from -I to +I by integer steps and can take the values for a nucleus with spin $\frac{1}{2}$ (e.g. ^{15}N, ^{13}C) of $m = +\frac{1}{2}$ and $m = -\frac{1}{2}$ respectively. These two permitted orientations, are termed as the α- and β-state, whereas α is the lower state ($m = +\frac{1}{2}$) and β the upper or excited state ($m = -\frac{1}{2}$).

However, in absence of a static magnetic field these two states are degenerate, this means that they have the same energy and no preferred orientation.

Table 1.1: Magnetic properties of some biologically useful nuclei

Nucleus	Spin quantum number I	Natural abundance [%]	Gyromagnetic ratio γ(10^7 $T^{-1}s^{-1}$)
^1H	1/2	99.99	26.7522
^2H	1	0.012	4.1066
^{14}N	1	99.63	1.934
^{15}N	1/2	0.37	-2.7126
^{13}C	1/2	1.07	6.7283
^{31}P	1/2	100	10.839

In a static magnetic field B_0, which by definition is oriented along the z-axis, the nuclear magnetic moment μ of the spins interact with the magnetic field resulting in a splitting of energy levels (Zeeman-splitting) (Figure 1.1, left panel). The energy E of the interaction between the magnetic moment μ_z (projection of μ on to B_0) and the magnetic field, is according to equation (Eq. 1.2),

$$E = -\mu B = -\mu_z B_0 = -\gamma I_z B_0 = -\gamma m \hbar B_0$$

<div align="right">**Eq. 1.2**</div>

where γ is the gyromagnetic ratio, a characteristic constant for a particular nucleus.

The equation Eq. 1.2 is of the utmost significance for the sensitivity of NMR experiments and illustrates that the energy difference ΔE is proportional to the magnetic field strength and depends on the gyromagnetic ratio γ of the corresponding nucleus. Table 1.1 summarizes relevant magnetic properties of some biologically useful nuclei. The energy difference between the two Zeeman energy states amounts to,

$$\Delta E = -\gamma \hbar B_0 = \hbar \omega_0$$

<div align="right">**Eq.1.3**</div>

where ω_0 is precessional frequency, or Lamor frequency of a nuclear spin, for an applied magnetic field strength. As described above, the two Zeeman energy states α and β will be unequally populated in a static magnetic field and the ratio can be predicted, at thermal equilibrium, by the Boltzmann distribution (Eq. 1.4),

$$\frac{N_\beta}{N_\alpha} = e^{-\Delta E/k_B T}$$

<div align="right">**Eq. 1.4**</div>

where k_B is the Boltzmann constant and T the temperature.

Since, the population differences for ^{1}H nuclei, which possess the highest gyromagnetic ratio among the spin $^1/_2$ nuclei in proteins, amount to only 0.002% for B_0 = 2.35 T at room temperature the sensitivity of NMR spectroscopy compared to other spectroscopic techniques is rather weak [29].

Moreover, the natural abundance of the NMR active nuclei ^{15}N and ^{13}C is very low, as can be seen in Table 1.1. To overcome the problem of detecting heteronuclei in biomacromolecules, special isotope enrichment techniques have been developed [32]. Furthermore, the availability of superconducting, high field magnets as well as of cryogenic circuitries and the development of relaxation optimized pulse sequences have improved the sensitivity of measurements in NMR spectroscopy [33].
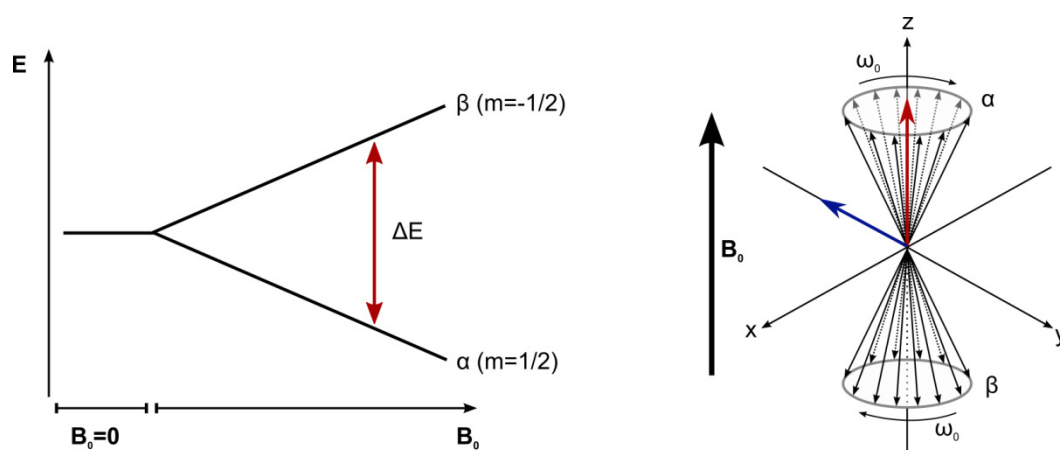
Figure 1.1: Basic concepts of NMR: (left panel): Energy level diagram: representation of the energy separation ΔE, between nuclear spin states α and β (Zeeman splitting) adopted by nuclei with spin $^1/_2$ without magnetic field (B_0=0) and with increasing field strength. **(right panel): Schematic illustration of the bulk magnetization in presence of the magnetic field:** Nuclear spins, represented by thin arrows, precess about the magnetic field along the z-axis of the laboratory frame according to their Lamor frequency ω_0. The macroscopic magnetization, labeled as red bolt arrow, is generated by the population differences between the two Zeeman states, and is parallel to the direction of the magnetic field B_0. Following a 90° radio frequency (rf) pulse along the x-axis, spins are phase coherent in the transverse plane, causing the projection of bulk transverse magnetization (blue bolt arrow) along the -y axis.

The sensitivity of NMR experiments is related to the signal to noise ratio (S/N) which can be described by the following equation (Eq. 1.5),

$$\frac{S}{N} \propto n\gamma_{exc}\gamma_{det}^{\frac{3}{2}}B_0^{\frac{3}{2}}(NS)^{\frac{1}{2}}T^{-1}T_2 \qquad \textbf{Eq. 1.5}$$

where n denotes the number of nuclear spins being observed, γ_{exc} the gyromagnetic ratio of the excited nuclei and γ_{det} that of the detected nuclei, B_0 the magnetic field strength, NS the number of accumulated spectra, T symbolize the temperature and T_2 the spin-spin relaxation rate, a time constant characterizing the signal decay. As implied by Eq. 1.5 the signal to noise ratio in NMR experiments cannot only be improved by higher magnetic field strengths but also by increasing the number of nuclei, which can be achieved by raising the sample concentration. Furthermore the S/N in NMR spectra depends on the quotient of the gyromagnetic ratios of the excited and of the detected nuclei. Most solution-state NMR experiments utilize ^{1}H excitation and detection in consequence of its high gyromagnetic ratio (see Table 1.1).

Based on the resonance condition (Eq. 1.2) transitions between the two Zeemann levels can be stimulated by electromagnetic radiofrequency (rf) pulses, according to equation Eq. 1.6,

$$\omega_{rf} = \omega_0 = \Delta E / \hbar$$

Eq. 1.6

where ω_{rf} is the angular frequency of the rf pulse, often called transmitter or carrier frequency and ω_0 is the Lamor frequency. The irradiation of a short rf pulse within the transverse plane, whose frequency is equal to the Lamor frequency of the respective spins (Eq. 1.6), results in phase coherence and the bulk magnetization vector is deflected in the xy- plane of the laboratory frame (Figure 1.1, right panel). This disturbance from the thermal equilibrium state, which induces an electromagnetic induction detectable by the coil, is the key to produce an NMR signal.

The observed variation in NMR signals for nuclear spins of the same isotope depends on the chemical environments of individual nuclei and they are referred to as chemical shift which is indicated by the symbol δ. This phenomenon is caused by an effect called *nuclear shielding* where the circulation of electrons generates local secondary magnetic fields which can augment or diminish the effect of the main field B_0. The effects of shielding on the precessional frequency ω of individual nuclei can be considered, according to Eq. 1.7,

$$\omega = -\gamma(1 - \sigma)B_0$$

Eq. 1.7

where σ is the average, isotropic, shielding constant.

Since, resonance frequencies are directly proportional to the static magnetic field B_0, the differences in chemical shift between two resonance signals measured in frequency units also rise or fall depending on the used magnetic field strength B_0.

In practice, chemical shifts are measured in *parts per million (ppm)* in terms of a standard reference, according to Eq. 1.8,

$$\delta = \frac{\Omega - \Omega_{\text{ref}}}{\omega_0} * 10^6 = (\sigma_{\text{ref}} - \sigma) * 10^6$$

Eq. 1.8

where Ω and Ω_{ref} are the offset frequencies of the signal of interest and the reference signal, respectively [27,28,30] The most common reference standards in biomolecular NMR are for ^{1}H and ^{13}C (DSS, TSP and TMS) and for ^{15}N (NH_3, NH_4Cl, urea and nitromethane) [34,35].

2.1.2 Relaxation

Relaxation is one of the fundamental aspects of magnetic resonance, which describes the coherence loss of the magnetization in the x-y plane as well as the return of the excited spins back to the equilibrium state. These two types of relaxation processes are defined as: The longitudinal or spin-lattice relaxation, which concerns the recovery of spin populations according to the Boltzmann equation and the transverse or spin-spin relaxation, which characterizes the decay of spin coherence. The longitudinal and transverse relaxation rates are termed R_1 and R_2 and their corresponding time constants are defined as T_1 and T_2, respectively. The behavior of the nuclear magnetization M as a function of time t in presence of relaxation times T_1 and T_2 can be described by the following equations (Eq. 1.9, Eq. 1.10)

$$M_z(t) = M_{z,eq}\left(1 - e^{-t/T_1}\right)$$

Eq. 1.9

$$M_{xy}(t) = M_{xy}(0)e^{-t/T_2}$$

Eq. 1.10

Time-dependent fluctuations of the magnetic or electric fields that fulfill the resonance condition (Eq. 1.2) and thus the stimulation of energy emission are the driving forces for the relaxation of excited spins back to the equilibrium. Fluctuations of local magnetic fields are generated by random molecular motions of other surrounding nuclei, changes in chemical shielding or unpaired electrons. In case of T_1 relaxation the exchange of energy with the surrounding ("lattice") occurs due to various intra- and intermolecular interactions, including dipolar-dipolar interactions, chemical shift anisotropy, spin rotation, scalar couplings and interactions with unpaired electrons in paramagnetic compounds.

The probability function of finding motions at a given angular frequency ω can be described by the spectral density function J(ω), according to Eq. 1.11,

$$J(\omega) = \frac{2\tau_c}{1 + (\omega\tau_c)^2}$$

Eq. 1.11

where τ_c represents the isotropic rotational correlation time. This characteristic time constant is expressed as the average time it takes for a molecule to move to a position at an angle of about 1 radian from its starting position and is used to describe the rate of random motions. For approximately spherical globular proteins τ_c can be calculated by Stoke's law (Eq. 1.12):

$$\tau_c = \frac{4\pi\eta r_H^3}{3k_B T} \text{ with } r_H = \sqrt[3]{\frac{3M_r}{4\pi\rho N_A}} + r_w$$

Eq. 1.12

in which η is the dynamic viscosity of the solvent, T the temperature, k_B is the Boltzmann constant and r_H is the effective hydrodynamic radius of the protein. The latter can be very roughly estimated from the molecular mass M_r of the protein, the Avogadro constant N_A as well as by assuming that the average density ρ for proteins is 1,37 g cm^{-3} and that a hydration layer of r_w = 1.6 to 3.2 Å surrounds the protein [28].

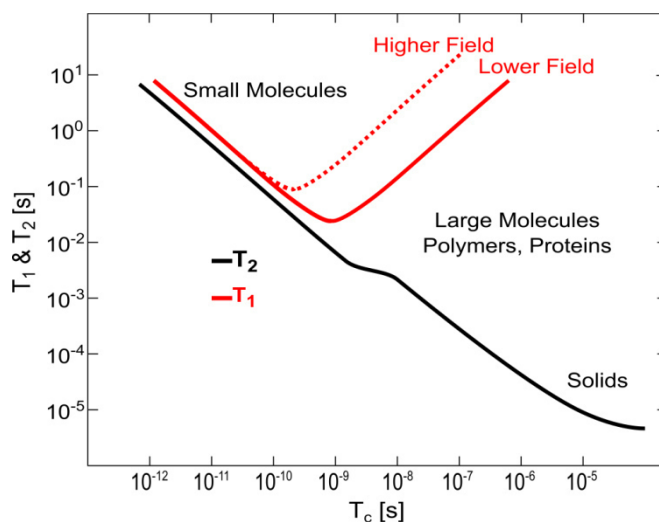**Figure 1.2: Schematically representation of the relaxation time constants T$_1$ and T$_2$ as a function of the rotational correlation time τ_c.** The diagram is based on the simplified theory in which all interactions are assumed to have the same correlation time. Within the figure T$_1$ is highlighted in red and T$_2$ is shown in black [adapted from [30,36,37]]

The equations Eq. 1.11 and Eq. 1.12 implies that the rate of molecular tumbling increases as a function of solvent viscosity and the hydrodynamic radius, which is related to the molecular size of the protein. For an aqueous solution of normal viscosity, T1 relaxation is inversely proportional to correlation time which is in the range of 10^{-12} to 10^{-13} seconds for small molecules with molecular weights less than 100, whereas macromolecules have a τ_c of around 10^{-8} seconds. A practical rule of thumb in estimating the rotational correlation time, is that the value of τ_c of a monomeric protein in an aqueous solution is approximately 0.6 times of its molecular weight in kDa. As shown in Figure 1.2. both the longitudinal and the transverse relaxation are related to the correlation time τ_c and it is evident that T_1 is severely affected by intermediate motions, and insensitive to very slow or very fast motions. T_2, however decreases with the same rate for small molecules or a slower rate for macromolecules (solids), hence $T_2 \leq T_1$. The slow tumbling of macromolecules (longer τ_c) in solution leads to faster relaxation of transverse magnetization (short T_2) due to enhanced spin-spin interactions and this is associated with an increased line width of NMR signals.

To overcome the problem of resonance broadening one simple, albeit limited, solution is to increase the overall molecular tumbling rate by recording NMR spectra at elevated temperatures. Another generally applicable approach is to suppress the spin diffusion by minimizing spin-spin interactions through perdeuteration of the proteins [38-40].

2.1.3 Nuclear Overhauser Effect (NOE)

The Nuclear Overhauser Effect (NOE) is defined as a change in the absorption intensity of one NMR resonance when the absorption of another resonance is perturbed from equilibrium by saturation or inversion. The nuclei involved may be either heteronuclear or chemically shifted homonuclear nuclei. The intensity change caused by NOE originates from zero-quantum and double-quantum transitions between the energy levels of spins that are dipolar coupled through space and are in spatial proximity to each other.
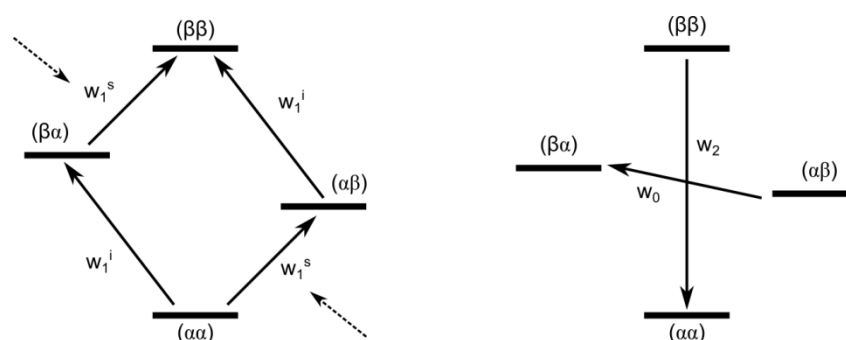

Figure 1.3: Relaxation pathways: Representation of the populations and energy levels (I - IV) for a two spin system with $J_{IS} = 0$ (left panel): The two "I" as well as the two "S" transitions have the same frequencies and are represented as w_1^I and w_1^S. The "S" transitions are saturated by continuous RF irradiation (shown as black dashed arrow) at the resonance frequency of spin "S". **Cross relaxation pathways immediately after saturation of the "S" transition (right panel):** The probabilities for zero-, and double-quantum transitions are represented by w_0 and w_2 [adapted from [30]].

The origin of the steady state NOE can be clearly illustrated for a two-spin system, in which the two spins "I" and "S" are coupled by dipolar-dipolar interactions under the assumption that no J coupling is between them (see Figure 1.3, left panel). By continuous irradiation of the "S" transitions for a sufficiently long time with a weak rf pulse, the target spin system is being saturated which results in equally populated Zeeman levels. After the irradiation, the system will try to restore the equilibrium through all allowable relaxation processes. Since the longitudinal relaxation pathways, denoted in the Figure 1.3 as w_1 do not alter the population differences of the spin s, the system relax via the dipole-dipole cross relaxation pathways w_0 and w_2 (see Figure 1.3, right panel).

It depends on the correlation time which of the pathways predominates the relaxation of the system, and hence the motion of the molecule. For small molecules (MW < 1 kDa) which are tumbling rapidly, double-quantum transitions (w_2) are the dominant relaxation pathway, resulting in observed intensity enhancement - positive NOE, whereas, zero-quantum transitions (w_0) are the prevailing relaxation pathway for large molecules (MW > 3 kDa) that are tumbling slowly.

This results in an observed intensity reduction and thus in a negative NOE. For medium sized molecules (MW 1 - 3 kDa) the NOE is very weak or may theoretically be zero, if w_2 = w0 (see Figure 1.4, left panel).

The NOE enhancement can be calculated by the steady state NOE enhancement factor η, according to the following equation (Eq. 1.13),

$$\eta = \frac{\gamma_S}{\gamma_I} \frac{(w_2 - w_0)}{(2w_1 + w_2 + w_0)}$$

Eq. 1.13

where $\gamma_{I/S}$ are the gyromagnetic ratios of the spin "I" and spin "S", respectively. The term (w_2 - w_0) represents the cross relaxation rate and the term (w_1 + w_2 + w_0) describes the total relaxation rate or dipolar spin-lattice relaxation. For homonuclear NOEs ($\gamma_I = \gamma_S$) the maximum enhancement is approximately fifty per cent (see Figure 1.4, right panel), whereas for heteronuclear NOEs the enhancement is also multiplied by the quotient of the gyromagnetic ratios.
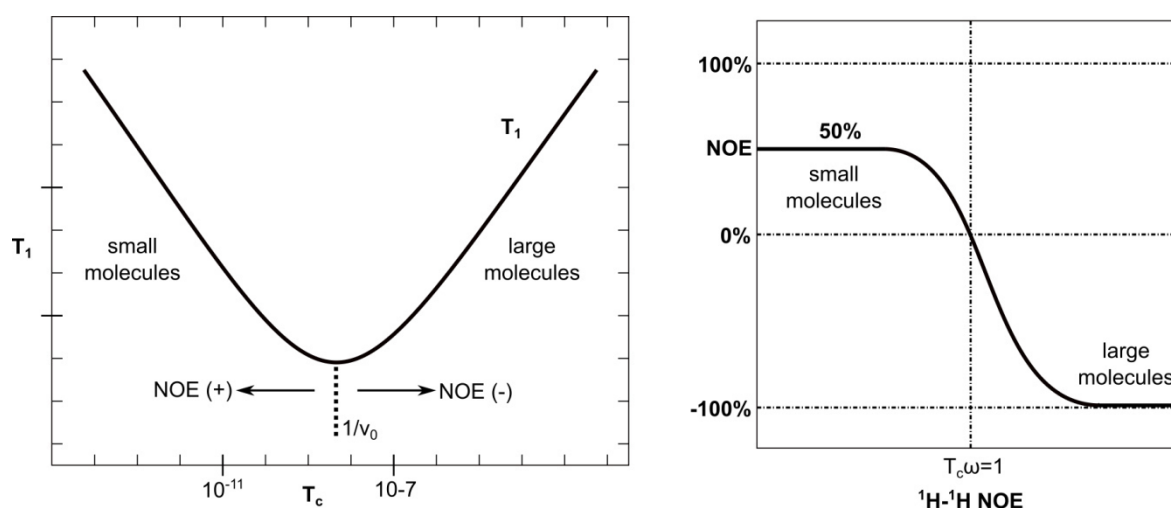
Figure 1.4: Graphic illustration of the effect of the molecular size and motion of molecules on the NOE: Representation of the longitudinal relaxation time constant T_1 as a function of the rotational correlation time τ_c. (left panel) Depending on the size of the molecule and / or the viscosity of the solvent, the NOE can be either positive or negative. **Representation of the signal intensity of homonuclear NOEs as a function of $\tau_c\omega$. (right panel):** The NOE enhancement is close to zero, when $\tau_c\omega$ ~ 1.

2.2 Two - Dimensional NMR experiments

2.2.1 Insensitive Nuclei Enhanced by Polarization Transfer (INEPT)

The INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) pulse sequence is the most utilized building block of many multidimensional NMR experiments which is used to transfer nuclear spin polarization from a sensitive nucleus with a high gyromagnetic ratio (usually ^{1}H) to a less sensitive nucleus with a lower gyromagnetic ratio (usually ^{15}N or ^{13}C). As a result, the detected signal from the heteronucleus will be increased. In contrast to NOE experiments, the INEPT step uses J-coupling for the polarization transfer and is thus independent of relaxation mechanisms. The pulse sequence as well as a vector representation of an standard INEPT experiment is shown in the figure below (see Figure 1.5).

The magnetization transfer of an INEPT sequence is as follows:

$$I\text{- spin:} \qquad 90°_x - \tau - 180°_x - \tau - 90°_y$$
$$S\text{ - spin:} \qquad 180°_x \qquad 90°_y - \text{acquisition}$$

The initial 90° excitation pulse on the I-spin (see Figure 1.5, right panel - step1) generates transverse magnetization (Eq. 1.14).

$$I_z + S_z = -I_y + S_z$$

Eq. 1.14

After one evolution delay of τ (with $\tau = 1/4J_{IS}$), simultaneous 180° pulses on both spins refocus the chemical shift but scalar couplings fully evolves for the whole of the period of 2τ, so that the spins undergo the following transformation (Eq. 1.15):

$$-I_y + (\frac{\gamma_S}{\gamma_I})S_z \xrightarrow{\pi I_x}\xrightarrow{\pi S_x}\xrightarrow{2\pi J_{IS}\tau 2I_z S_z} I_y \cos 2\pi J_{IS}\tau - 2I_x S_z \sin 2\pi J_{IS}\tau - (\frac{\gamma_S}{\gamma_I})S_z$$
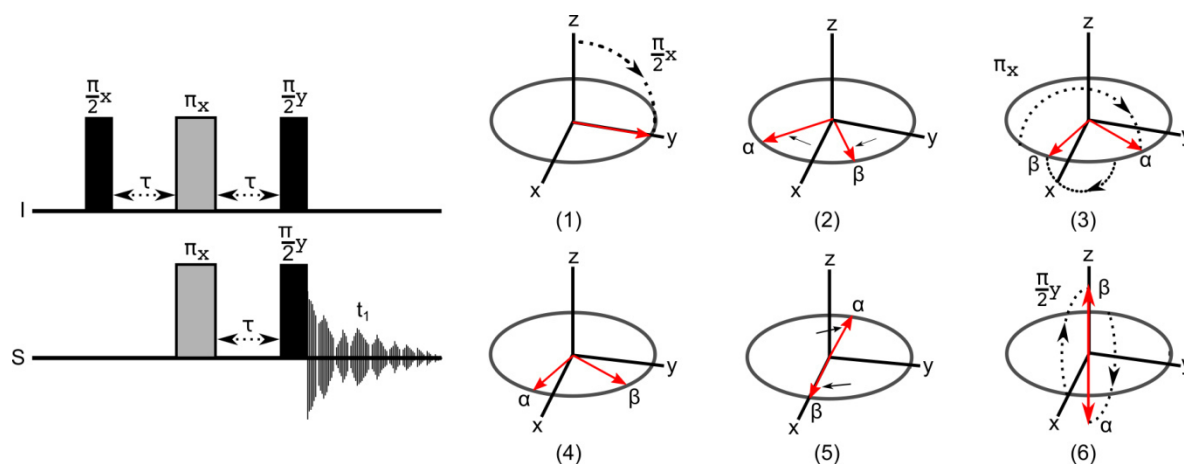
Eq. 1.15

Figure 1.5: The basic INEPT experiment: Pulse sequence (left panel): The enhancement of the NMR signals results from the transfer of a nuclear spin polarization from the I spin, with large Boltzmann population differences such as ^{1}H, to the nuclear spin of interest (S spin) usually an insensitive nuclei with low Boltzmann population differences such as ^{15}N or ^{13}C. The fixed delay τ is equal to $1/4J_{IS}$, 90° pulses are shown as filled black rectangles and 180° pulses are highlighted in grey. **Behavior of the I spin magnetization vectors in an INEPT experiment (right panel):** After an initial 90°$_x$ pulse the two magnetization vectors are flipped in the xy plane (1) where they process for a period of τ until they accumulate a relative phase angle of 90° (2). A 180° pulse applied on the I spins inverts both vectors within the transverse plane (3) while at the same time a second 180° pulse on the S spins interchanges the I spin labels α and β (4). Free precession for a further period of τ leads to an antiparallel formation of the vectors along the ± x-axis (5). The last 90° pulse about the y-axis realigns the magnetization vectors to the longitudinal direction (6) [adapted from [41]].

The following 90° pulse rotates the I-spin magnetization components back to the z axes and the 90° pulse on the S-spin brings the coupled magnetization components back to the x-y plane, which results in detectable anti-phase magnetization (Eq. 1.16, Eq. 1.17).

$$I_y \cos 2\pi J_{IS}\tau \pm 2I_z S_z \sin 2\pi J_{IS}\tau - \left(\frac{\gamma_S}{\gamma_I}\right)S_z \qquad \textbf{Eq. 1.16}$$

$$I_y \cos 2\pi J_{IS}\tau \pm 2I_z S_y \sin 2\pi J_{IS}\tau - \left(\frac{\gamma_S}{\gamma_I}\right)S_y \qquad \textbf{Eq. 1.17}$$

After the subtraction of the two experiments with an alternating phase, the following term is retained (Eq. 1.18).

$$-2I_z S_y \sin 2\pi J_{IS}\tau \qquad \textbf{Eq. 1.18}$$

Through the INEPT sequence, the sensitivity of the S-spin is enhanced by spin I for a factor of γ_I / γ_S and the signal maximum will be achieved, when the term $\sin 2\pi J_{IS}\tau = 1$.

2.2.2 The HSQC experiments

The HSQC (heteronuclear single-quantum coherence) and HMQC (heteronuclear multiquantum correlation) are the basic building blocks of many multidimensional experiments, described in the next section. The HSQC experiment employs two INEPT magnetization transfers. The first one creates an antiphase heteronuclear coherence whereas the second one is used to convert this coherence back to an observable magnetization. The enhancement in sensitivity which is reached by this sequence is significantly higher than the enhancement obtained by using the nuclear Overhauser effect in simple heteronuclear correlated spectra. The pulse sequence of a ^{1}H-^{15}N HSQC experiment is outlined in figure 4.

The first INEPT step (see Figure 1.5) generates a proton antiphase magnetization during the period τ. To refocus chemical shift modulation during this period, a 180°C pulse on both spins will be applied in the middle of these period. Subsequently, the coherence is transferred to the directly attached heteronucleus (usually ^{15}N or ^{13}C) by the simultaneous application of 90° pulses on both spins.

The antiphase heteronuclear single quantum coherence evolves during the subsequent t_1 evolution period and a 180° pulse in the middle of these period refocuses the evolution of the ^{1}H heteronuclear J_{IS} scalar coupling interaction.

The second INEPT step is used to transfer the frequently labeled heteronuclear single quantum coherence back to the ^{1}H magnetization for detection. Evolution in terms of the product operators is described for a two-spin system as follows (Eq. 1.19).

$$I_z \xrightarrow{I_x} -I_Y \xrightarrow{2I_{IS}I_zS_z} -I_Y \cos \pi J_{IS}\tau + 2I_xS_z \sin \pi J_{IS}\tau \xrightarrow{\tau=1/2J_{IS}} 2I_xS_z \xrightarrow{(I_{-Y}+S_x)} -2I_zS_Y$$

$$\xrightarrow{I_y \,\&\, (\Omega_s t_1)S_z} 2I_zS_Y \cos \Omega_s t_1 - 2I_zS_x \cos \Omega_s t_1 \xrightarrow{(I_Y+S_x)} 2I_xS_z \cos \Omega_s t_1 \qquad \textbf{Eq. 1.19}$$

$$-2I_xS_x \sin \Omega_s t_1 \xrightarrow{2I_{IS}I_zS_z} I_Y \cos \Omega_s t_1 - 2I_xS_x \sin \Omega_s t_1$$

where I refers to proton spin and *S* to the heteronucleus spin. The delay τ is given by $\tau = 1/4J_{IS}$ and the chemical shift of the heteronucleus is denoted as Ω_S. Prior to the detection there is the in-phase magnetization represented by the operator $I_Y \cos \Omega_s t_1$ as well as a multiple quantum coherence term $2I_xS_x$.

The second one is unobservable and can therefore be neglected. Moreover, two-step phase cycles are necessary to filter out magnetization obtained from protons attached to ^{12}C or ^{14}N nuclei and four step phase cycles for the complete elimination of undesirable double quantum terms.
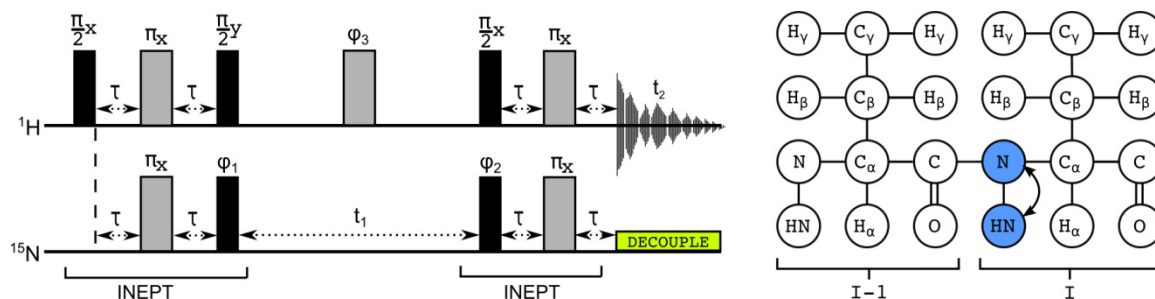

Figure 1.6: ^{1}H-^{15}N HSQC experiment - Representation of the pulse sequence (left panel): Filled black rectangles and filled grey rectangles represent 90° and 180°, respectively. The nominal value for $2\tau = 1/(2J_{NH})$. Phase cycling is given by: Φ_1 = x, -x; Φ_2 = 2(x), 2(-x); Φ_3 = 4(y), 4(-y) and the receiver = 2(x, -x, -x, x) [adapted from [28]]. **Schematic representation of the correlations occurring in an HSQC within the protein backbone (right panel)**.

Additionally to the conventional HSQC pulse sequence, described above (see example for ^{1}H-^{15}N HSQC in Figure 1.6), a number of modifications have been proposed in the literature to improve the sensitivity, to achieve better resolutions or to remove peak artifacts. Pulse field gradient (PFG) versions for example were incorporated to suppress undesirable coherence pathways and to achieve a superior suppression of solvent signals. Moreover, PFG's allow a clear distinction between ^{1}H - ^{12}C versus ^{1}H - ^{13}C magnetization, which results in the collection of high quality HSQC spectra under standard routine conditions [42,43]. The Echo / Anti-Echo method is one of the most widely used protocols to achieve coherence selection by using PFG's [44]. On the other hand the preservation of equivalent pathways (PEP) methodology was introduced to enhance the sensitivity by the factor of $\sqrt{2}$ for a two spin system through the implementation of a second refocused INEPT step in a regular HSQC pulse sequence [45]. Moreover, in 1992 Kay and coworkers have published a sensitivity-enhanced HSQC pulse sequence which combines the sensitivity enhancement of the PEP methodology with pulse field gradients for coherence selection [45]. This combination was implemented later on for a three dimensional HNCO [46,47]. Nowadays, both modifications are frequently used to improve a variety of multidimensional NMR experiments [48].

2.3 Triple Resonance Experiments

2.3.1 HNCA and HNCO

The HNCA experiment correlates ^{15}N and NH chemical shifts with interresidue ^{13}C$^{\alpha}$ shifts (see Figure 1.8, right panel), whereas the HNCO experiment correlates the amide ^{1}H and ^{15}N chemical shifts with the ^{13}C shifts of the carbonyl resonance of the preceding amino acid (see Figure 1.7, right panel). The figures Figure 1.7 (left panel) and Figure 1.8 (left panel) represents the pulse sequences of an HNCO and a non constant time HNCA, respectively. In both experiments the magnetization originating from the amide ^{1}H is transferred to the attached ^{15}N via J_{NH} coupling during INEPT sequence. In the subsequent t_1 period, the 15N magnetization evolves exclusively under the influence of ^{15}N chemical shift whereas ^{1}H, ^{13}C', and ^{13}C$_{\alpha}$ decoupling is achieved by applying 180° refocusing pulses in the middle of t_1 period. After the evolution of t_1, the ^{15}N magnetization becomes antiphase during the delay δ with respect to the coupled ^{13}C$_{\alpha}$ spins in the HNCA and with respect to the polarization of the carbonyl spin of the preceding residue in the HNCO.

Figure 1.7: 3D HNCO experiment - Representation of the pulse scheme (left panel): Filled black rectangles and filled grey rectangles represent 90° and 180°pulses, respectively. Pre-saturation of the water resonance is required during the recovery period and during both δ periods. The nominal value for $2\tau = 1/(2J_{NH})$. The phase cycling scheme employed is as follows: $\Phi_1 = $ x, -x; $\Phi_2 = $ 2(y), 2(-y); $\Phi_3 = $ x; $\Phi_4 = $4(x), 4(-x); $\Phi_5 = $8(x), 8(y), 8(-x), 8(-y) and the receiver is given by = 2(x), 4(-x), 2(x), 2(-x), 4(x), 2(-x) [adapted from [28]].**Schematic representation of the correlations occurring in an HNCO within the protein backbone (right panel).**

Thereafter, antiphase ^{15}N magnetization will be converted into ^{15}N - ^{13}C' two spin coherence (HNCO) or in NH - ^{15}N - ^{13}C$_{\alpha}$ three spin coherence (HNCA) by applying a 90° pulse to the ^{13}C' spins (HNCO - 90° pulse (Φ_4)) or to both ^{1}H and ^{13}C$_{\alpha}$ spins (HNCA, see Figure 1.8).

The contributions of ^{15}N and ^{1}H chemical shifts are removed during the t_2 period by the execution of 180° pulses. In addition, ^{13}C' - ^{13}C$_\alpha$ decoupling is achieved by applying 180° pulses at the midpoint of t_2 to the ^{13}C$_\alpha$ (HNCO) or to the ^{13}C' (HNCA) respectively. Finally, the magnetization is transferred back to the amide protons by reversing the transfer steps described previously, with the exception that the t_1 period is omitted.
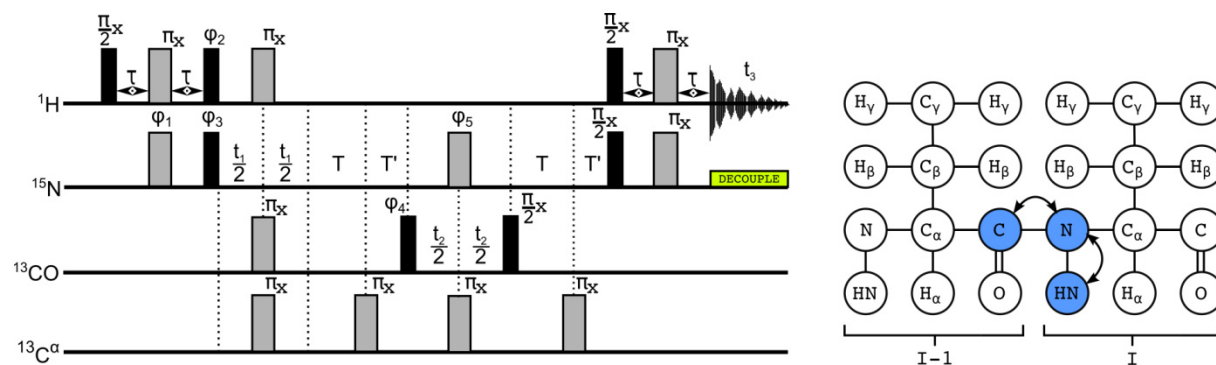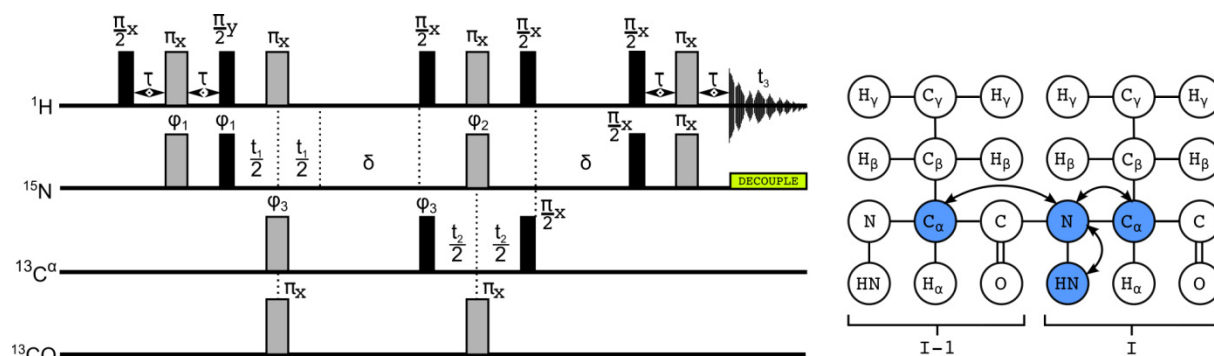

Figure 1.8 Non constant time 3D HNCA experiment - Representation of the pulse sequence (left panel): Filled black rectangles and filled grey rectangles represent 90° and 180°pulses, respectively. The delay δ is adjusted to be an integral multiple of $1/J_{NH}$ and to allow maximal magnetization transfer between the ^{15}N and ^{13}C spins. The phase cycling employed is as follows: Φ_1 = x, -x; Φ_2 = 4(x), 4(y), 4(-x), 4(-y); Φ_3 = 2(x), 2(-x) and the receiver is given by = x, -x, -x, x, -x, x, x, -x [adapted from [28]] **Schematic representation of the correlations occurring in a HNCA within the protein backbone (right panel).**

The HNCA pulse sequences can be described concisely by the operator formalism as,

$$I_z \longrightarrow -2I_zN_y \longrightarrow 2I_zN_y \cos\Omega_N t_1 \cos\pi J_{NC'}t_1 \longrightarrow -4I_yN_xA_y \cos\Omega_N t_1 \cos\pi J_{NC'}t_1$$

$$\longrightarrow 4I_yN_xA_y \cos\Omega_N t_1 \cos\pi J_{NC}t_1 \cos\Omega_A t_2 \qquad \textbf{Eq. 1.20}$$

$$\longrightarrow I_x \cos\Omega_N t_1 \cos\pi J_{NC'}t_1 \cos\Omega_A t_2$$

under the assumption that the two-bond coupling between ^{15}N(i) and ^{13}C$_\alpha$(i-1) is zero and that the delay δ=1/(2J_{NA}). The characters I, N, and A denote the intraresidue ^{1}H, ^{15}N and ^{13}C spins and the carbonyl spin of the preceding residue is designated by C'. The chemical shifts of the intraresidue ^{15}N and C$_\alpha$ are denoted as Ω_N and Ω_A and $J_{NC'}$ represents the interresidue ^{15}N-C' coupling constant. Eq. 1.20 indicates that the detected signal is amplitude modulated by the t_1 and t_2 dimensions, resulting in pure absorption lineshapes after a three dimensional Fouríer Transformation. When recording a HNCA experiment two sets of correlations are obtained: the intraresidue C$_\alpha$(i) - ^{15}N(i) - NH(i) connectivity via one bond $J_{NC\alpha}$ coupling and the interresidue C$_\alpha$(i-1) - ^{15}N(i) - NH(i) connectivity via two bond $J_{NC\alpha}$ coupling (see also Figure 1.8).

Denoting ^{1}H, ^{15}N and ^{13}C spins by I, N, S, respectively, the pulse sequence of the HNCO experiment can be described as follows,

$$I_z \longrightarrow -2I_zN_y \xrightarrow{t_1} 2I_zN_y \cos\Omega_N t_1$$

$$\longrightarrow 4N_xS_yI_z \cos\Omega_N t_1$$

$$\xrightarrow{t_2} 4N_xS_yI_z \cos\Omega_N t_1 \cos\Omega_S t_2 \cos\pi J_{NC_\alpha} t_2 \left\{ \prod_k \cos\pi J_{ks} t_2 \right\} \qquad \text{Eq. 1.21}$$

$$\longrightarrow I_x \cos\Omega_N t_1 \cos\Omega_S t_2 \cos\pi J_{NC_\alpha} t_2 \left\{ \prod_k \cos\pi J_{ks} t_2 \right\}$$

assuming that the J coupling and the chemical shift effects, which are removed by 180° pulses, have been omitted and that T=T'. Moreover, the two bond interresidue ^{15}N - C_α is often small in proteins and has been neglected in Eq. 1.21. Ω_N and Ω_S denote the 15N and C' chemical shift offsets and $J_{NC\alpha}$ refers to the intraresidue ^{15}N-C_α coupling and J_{KS} to the coupling between the carbonyl carbon and other protons, k. Eq. 1.21 shows that the detected in-phase signal (I_x) is modulated in amplitude by the chemical shift offsets Ω_N and Ω_S in the t_1 and t_2 dimension, respectively, evoking a purely absorptive 3D NMR signal after three dimensional Fourier Transformation.

The highly sensitive 3D HNCO and the 3D HNCA experiment are powerful tools in the sequential assignment of NMR resonances in protein backbones, but no information can be obtained on the side chain atoms of the individual amino acids.

2.3.2 CBCANH / HNCACB

The CBCANH [49] and the HNCACB [50] are standard tools in biomolecular NMR to correlate the resonances of the amide moiety with alpha and beta carbon resonances in proteins, whether these are intra- or interresidue correlations (Figure 1.9 - right panel). However, the HNCACB has some advantages over the CBCANH experiment. Thus it could be shown that the HNCACB is more sensitive, because it exploits the long T_2 relaxation properties of the amide ^{15}N nuclei relative to those of the $^{13}C_{\alpha/\beta}$ spins. This should provide higher S/N ratios for large proteins where sensitivity losses due to relaxation become problematic. Moreover, the HNCACB experiment can be used to assign the side chain amide groups and carbon atoms of the amino acids asparagines and glutamine. The pulse sequence of a commonly used out and back HNCACB is outlined in Figure 1.9 (left panel). It represents an extension of the HNCA experiment originally described by Kay et al [51] and contains several well-known pulse sequence building blocks. The magnetization transfer of an HNCACB experiment is as follows:

$$H^N \longrightarrow N \longrightarrow C_\alpha \longrightarrow C_{\alpha,\beta}(t_1) \longrightarrow C_\alpha \longrightarrow N(t_2) \longrightarrow H^N(t_3)$$
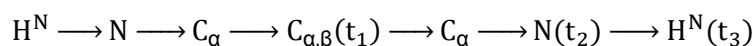
After an initial INEPT transfer from ^1H to ^{15}N, the ^5N magnetization is allowed to refocus relative to the coupling to the H^N for a period $\delta_1 = 1/(2J_{NH})$, after which the proton decoupling is turned on. Throughout the period of the half of the delay τ, a buildup of antiphase ^{15}N magnetization, relative to the $^{13}C_\alpha$ resonances of the intra residue (i) and of the preceding residue (i-1), occurs. Thereafter, antiphase ^{15}N magnetization is converted to antiphase coherence with respect to the $^{13}C_\alpha$ spins by the simultaneous application of 90° pulses on nitrogen and carbon (Φ_2). During the following period of $2\delta_2$, the $J_{C\alpha C\beta}$ scalar coupling is active, due to the fact that the 180° pulse applied to $^{13}C_\alpha$ / $^{13}C_\beta$ in the middle of this period is non-selective. Following the first period of $2\delta_2$, the relevant terms are given by Eq. 1.22:

$$-N_z C_y^\alpha (\cos \pi J_{AB}\{2\delta_2\}) + 2N_z C_x^\alpha C_z^\beta (\sin J_{AB}\{2\delta_2\}) \qquad \textbf{Eq. 1.22}$$

The relative magnitudes of the transfer functions are determined by the value of $2\delta_2$. If $2\delta_2 = 1/(8\ J_{C\alpha C\beta})$ the magnitudes of the operators are approximately equal and correlations to $^{13}C_\alpha$ and $^{13}C_\beta$ are observed in the final spectra, whereas only $^{13}C_\beta$ are observed in the final spectra with a maximum magnitude, if the value of $2\delta_2 = 1/(4\ J_{C\alpha C\beta})$. Subsequently, a 90° pulse (Φ_3) rotates the carbon magnetization.

$$-N_z C_y^\alpha - 2N_z C_z^\alpha C_x^\beta \qquad\qquad \textbf{Eq. 1.23}$$

During t_1 evolution time, the chemical shifts of the $^{13}C_\alpha$ and $^{13}C_\beta$ nuclei as well as the carbon - carbon scalar couplings ($J_{C\alpha C\beta}$, $J_{C\beta C\gamma}$) evolve, whereas nitrogen and carbonyl spins are decoupled by 180° pulses.
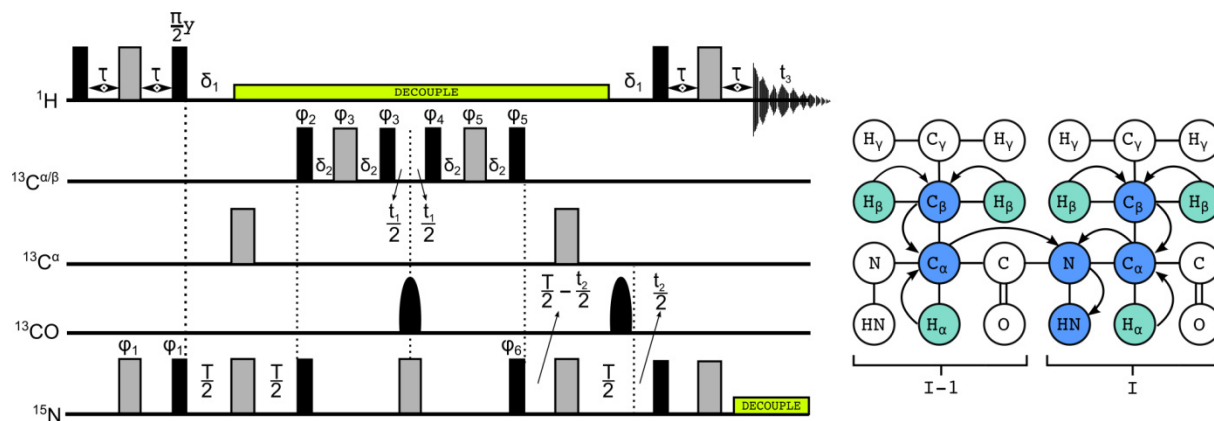
Figure 1.9: 3D HNCACB experiment - Representation of the pulse scheme of an out and back (left panel): Filled black rectangles and filled grey rectangles represent 90° and 180° pulses, respectively with phase x, unless indicated. Rounded bars represent selective 180° pulses applied to ^{13}CO spins. The phase cycle is given by Φ_1 = x, -x; Φ_2 = 2(x), 2(-x); Φ_3 = y, Φ_4= y, Φ_5= x, Φ_6= 4(x), 4(-x) and the receiver by = x, -x, x, -x, x, x, -x. States TPPI is employed in t_1 and t_2 to obtain quadrature detection and the phase rotation, induced by the non-resonant effects of the 180° pulses applied to the ^{13}CO spins during t_1 can be compensated by adding a correction to pulse phases Φ_4 and Φ_5 [adapted from [28]]. **Schematic representation of the correlations occurring in an HNCACB within the protein backbone (right panel)**

After the following 90° pulse the magnetization can be described according to,

$$-N_z C_y^\alpha (\cos \pi\, \Omega_{C\alpha} t_1) + 2N_z C_x^\alpha C_z^\beta (\sin \Omega_{C\beta}\, t_1) \qquad\qquad \textbf{Eq. 1.24}$$

where $\Omega_{C\alpha}$ and $\Omega_{C\beta}$ are the chemical shifts of the corresponding spins. After refocusing of the preceding operators during the second period of $2\delta_2$ the magnetization can be described by,

$$-N_z C_y^\alpha (\cos \pi\, \Omega_{C\alpha} t_1) + 2N_z C_y^\alpha (\sin \Omega_{C\beta}\, t_1) \qquad\qquad \textbf{Eq. 1.25}$$

The terms for the $^{13}C_\alpha$ and $^{13}C_\beta$ chemical shifts are of sign opposite, indicating that the correlation peaks for the corresponding $^{13}C_\alpha$ and $^{13}C_\beta$ spins are of opposite phase.

The inverted phases are useful for distinguishing of arising crosspeaks in the final spectra for both manual and automated assignment. Within the remaining pulse sequence the same magnetization transfer steps are repeated in opposite order to transfer the magnetization back to the H^N, for detection [28,50,52].

2.4 Nuclear Overhauser Enhancement Spectroscopy (NOESY)

All of the triple resonance experiments described above, based on magnetization or coherence transfer via scalar coupling to provide orientational information between protons and heteronuclei of the same or the preceding residue within a protein backbone [28]. Therefore, no information on the distance between spins for structural and dynamic studies of molecules in solution can be obtained by J - modulated experiments. However, NOESY experiments provide short range as well as long range internuclear distance restraints by using dipolar couplings in the form of cross relaxation, to correlate protons that are close in space (distance smaller than 5 Å). This makes NOESY experiments an attractive tool in the structural characterization of molecules in solution. Small size molecules, are generally measured using one dimensional NOESY experiments, whereas for larger molecules such as proteins two or multidimensional experiments are preferred, especially to overcome signal overlaps.

2.4.1 3D ^{15}N edited HSQC NOESY

The 3D ^{1}H-^{15}N NOESY HSQC experiment is a combination of a standard NOESY sequence after removing the acquisition period with a standard HSQC sequence [53,54]. The pulse sequence of such a ^{1}H-^{15}N NOESY HSQC experiment is illustrated in Figure 1.10 and the magnetization transfer is given by,

$$H(t_1) \xrightarrow{\text{NOE}} H \xrightarrow{J_{HN}} N(t_2) \xrightarrow{J_{HN}} H(t_3)$$
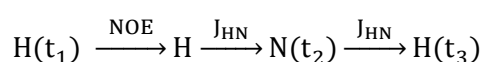
Initially, an element consisting of the t_1 evolution period flanked by two 90° proton pulses ($90°_{\Phi 1}$ - t_1 - $90°_{\Phi 2}$), is used to frequency label the spins and returns the magnetization back to the z-axis. The scalar coupling between ^{15}N and ^{1}H is refocused by a 180° pulse to the heteronuclei at the midpoint of the t_1 evolution time.

Following the evolution of the ^{1}H chemical shifts, the magnetization is transferred to vicinal protons via cross relaxation during the NOESY mixing time τ_m. The first 90° proton pulse following τ_m is equivalent to the first 90° proton pulse in a HSQC experiment [28,31]. Afterwards, the magnetization of the H^N spins is transferred to ^{15}N and then back to H^N after the evolution of ^{15}N chemical shifts in the same pathway as in a standard HSQC experiment [31].

For two ^{1}H spins (H^A and H^B) and a ^{15}N spin that is scalar coupled only to the ^{1}H$_A$ spin, the coherence transfer through ^{1}H-^{15}N NOESY HSQC pulse sequence is given concisely by the operator formalism as [28],

$$
H_z^A + H_z^B \xrightarrow{\frac{\pi}{2}(H_x^A+H_x^B)-\frac{t_1}{2}-\pi N_x-\frac{t_1}{2}-\frac{\pi}{2}(H_x^A+H_x^B)} H_z^A \cos(\Omega_{H^A}t_1)\prod_\eta \cos(\pi J_{H^A_\eta}t_1)
$$

$$
+ H_z^B \cos(\Omega_{H^B}t_1)\prod_\lambda \cos(\pi J_{H^B_\lambda}t_1)
$$

$$
\xrightarrow{\tau_m} H_z^A[\alpha_{H^A H^A}(\tau_m)\cos(\Omega_{H^A}t_1)\prod_\eta \cos(\pi J_{H^A_\eta}t_1)
$$

$$
+\alpha_{H^A H^B}(\tau_m)\cos(\Omega_{H^B}t_1)\prod_\lambda \cos(\pi J_{H^B_\lambda}t_1)]
$$

$$
\xrightarrow{\frac{\pi}{2}(H_x^A+H_x^B)-\tau-\pi(H_x^A+H_x^B),\pi N_x-\tau-\frac{\pi}{2}(H_y^A+H_y^B),\frac{\pi}{2}(N_x)}
$$

$$
-2H_z^A N_y[\alpha_{H^A H^A}(\tau_m)\cos(\Omega_{H^A}t_1)\prod_\eta \cos(\pi J_{H^A_\eta}t_1)
$$

$$
+\alpha_{H^A H^B}(\tau_m)\cos(\Omega_{H^B}t_1)\prod_\lambda \cos(\pi J_{H^B_\lambda}t_1)]\xrightarrow{\frac{t_2}{2}-\pi(H_x^A+H_x^B)-\frac{t_2}{2}}
$$

$$
-2H_z^A N_y \cos(\Omega_N t_2)[\alpha_{H^A H^A}(\tau_m)\cos(\Omega_{H^A}t_1)\prod_\eta \cos(\pi J_{H^A_\eta}t_1)
$$

$$
+\alpha_{H^A H^B}(\tau_m)\cos(\Omega_{H^B}t_1)\prod_\lambda \cos(\pi_{H^B_\lambda}t_1)]
$$

$$
\xrightarrow{\frac{\pi}{2}(H_x^A+H_x^B),\frac{\pi}{2}N_x-\tau-\pi(H_x^A+H_x^B),\pi N_x-\tau}
$$

$$
-H_x^A \cos(\Omega_N t_2)[\alpha_{H^A H^A}(\tau_m)\cos(\Omega_{H^A}t_1)\prod_\eta \cos(\pi J_{H^A_\eta}t_1)
$$

$$
+\alpha_{H^A H^B}(\tau_m)\cos(\Omega_{H^B}t_1)\prod_\lambda \cos(\pi J_{H^B_\lambda}t_1)]
$$

Eq. 1.26

in which the evolution time is given by $\tau = 1/(4J_{HN})$, Ω represents the chemical shift of the corresponding spin and $\alpha_{H^A H^A}(\tau_m)$ and $\alpha_{H^A H^B}(\tau_m)$ are transfer functions for dipolar cross - relaxation. Homonuclear scalar couplings to other proton spins are represented by the terms $J_{H^A_\eta}$ and $J_{H^B_\lambda}$.
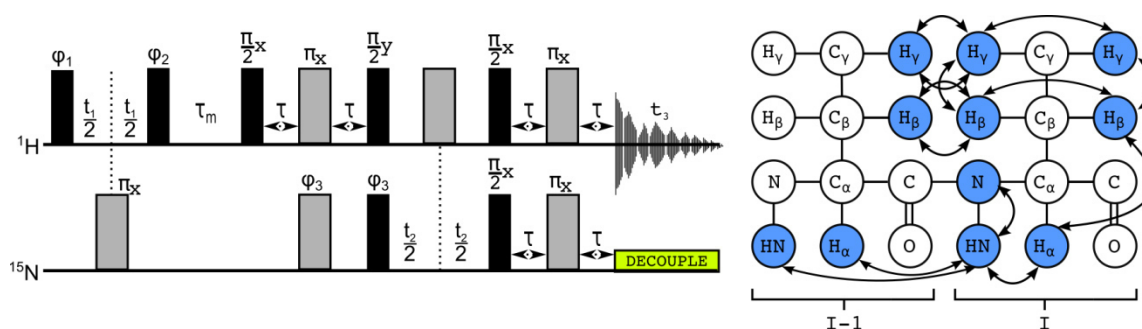
Figure 1.10: 3D ^{1}H - ^{15}N NOESY HSQC experiment - Representation of the pulse scheme (left panel): Filled black rectangles and filled grey rectangles represent 90° and 180° flip angles, respectively. The nominal value for $\tau=1/(4J_{HN})$ and τ_m is the mixing time. The phase cycle is given by $\Phi_1 = 2(x), 2(-x); \Phi_2 = 4(x), 4(-x); \Phi_3 = x, -x;$ and the receiver by x, -x, -x, x, -x, x, x, -x (adapted from [28]) **Schematic representation of the correlations occurring in an ^{1}H - ^{15}N NOESY HSQC within the protein backbone (right panel).**

As mentioned above, correlations between spins that are close in space occur via cross relaxation. The intensity of the arising cross peaks, is not only proportional to the cross relaxation rate, but also comes from spin diffusion and increases as the mixing time τ_m increases. However, spin diffusion can be minimized, by using shorter mixing times. Usually, a mixing time of 30 - 200 ms is recommended for measurements of protein samples to avoid the effect of spin diffusion. Another limitation of NOESY experiments exists for the measurement of small and medium size molecules (molecular weight app. 1000 Da) in the "crossover" region where $\omega\tau_c$ is approximately 1, resulting in an NOE enhancement close to zero (see Figure 1.4 - right panel) [31].

2.5 Residual Dipolar Couplings (RDC)

In recent years Residual Dipolar Couplings (RDC) have emerged as a new tool for analyzing of biomolecules by liquid-state NMR. They provide crucial global orientational information and significantly complements classical short-range restraints obtained from the chemical shift, the scalar coupling constant as well as the NOE's in the determination of biomolecular structures [55]. Moreover, RDC's can provide valuable information regarding internal dynamics over a wide range of timescales [56,57]. The structural and dynamic insights gained from measurements of RDC's make them extremely useful for investigations of protein folding, ligand - protein and protein - protein interactions. Furthermore, they will be used for the structure determination of nucleic acids as well as for the analysis of carbohydrates in its free form or in protein-bound state [58,59].

For a more complete overview of the applications of RDC's in structural biology and for the physical theory behind , the reader is being referred to the following reviews [55,56,58,60-64].

2.5.1 Theoretical Framework

The physical basis of residual dipolar coupling's (RDC) is the dipole - dipole (DD) interaction between two magnetic active nuclei in an external magnetic field B_0. As a result of rotational Brownian diffusion, dipolar couplings average to zero under isotropic solution conditions but can be observed under slightly anisotropic solution or solid-state conditions. The spin-spin coupling for the nuclei I and S in the anisotropic phase is given by (Eq. 1.27),

$$\omega_{IS} = D_{IS} + \pi J_{IS} \qquad \text{Eq. 1.27}$$

where the secular part of the dipole-dipole coupling D_{IS} can be described by the following equation (Eq. 1.28),

$$D_{IS} = D_{max}\langle(3\cos^2\theta - 1)/2\rangle \qquad \text{Eq. 1.28}$$

where θ is the angle between the internuclear bond vector and the external magnetic field B_0 (see Figure 1.11, left panel). The pointed brackets around the θ dependent term denote time or ensemble averaging and D_{max} is the dipole-dipole coupling constant, which can be determined according to Eq. 1.29,

$$D_{max} = -\frac{\mu_0 h \gamma_A \gamma_B}{8\pi^3 r_{AB}^3}$$

Eq. 1.29

where μ_0 is the vacuum permeability, h is the Planck's constant, γ_I and γ_S are the gyromagnetic ratios of nuclei I and S, respectively, and r_{IS} is the distance between nuclei I and S. The dipole - dipole coupling constant D_{max} is bond type dependent and usually on the order of 10^3 Hz (see Figure 1.11, middle panel). Furthermore, it can be seen from equation Eq. 1.29 that the value of D_{max} strongly depends on the internuclear distance between the coupled nuclei I and S, which means that D_{max} decreases rapidly with increasing internuclear distance as a function of r^{-3}.
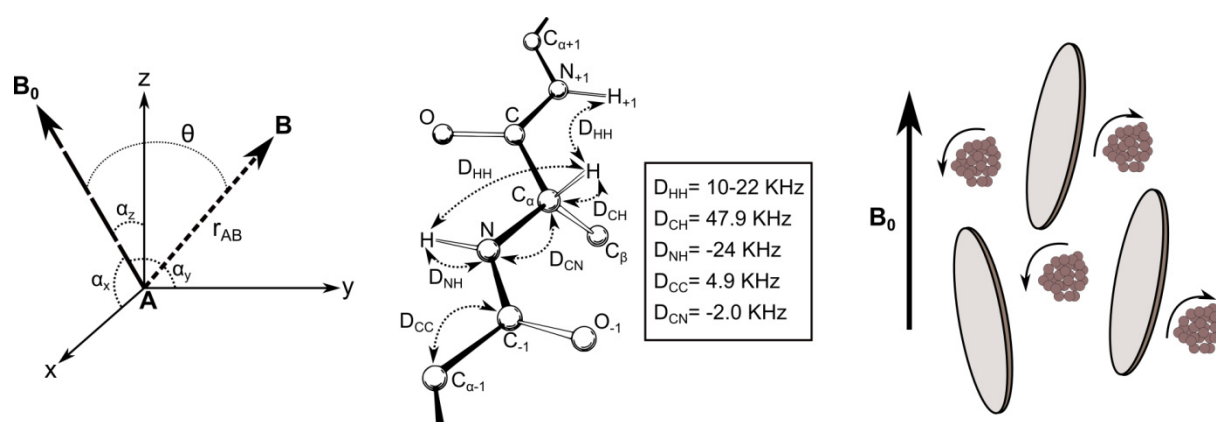
Figure 1.11: Basics of RDC's - Diagram illustrating the relationships between an RDC internuclear vector IS and an arbitrary molecular frame (left panel): **B$_0$** represents the external magnetic field and θ is the instantaneous angle between the internuclear vector IS and B$_0$. The projection angles of B$_0$ onto each axis of the molecular frame are given by α$_x$, α$_y$, α$_z$. **r$_{IS}$** represents the distance between the nuclei I and S. **Schematic representation of characteristic values for the static dipolar coupling constant in a protein backbone [64] (middle panel). Illustration of the steric alignment of biomolecules with bicelles aligning with their bilayer normal perpendicular to the external magnetic field B$_0$ (right panel).**

2.5.2 Measurement of residual dipolar couplings

In recent years, there have been numerous experiments designed to accurately measure RDC's and they are all based on two basic techniques . The frequency displacement technique, where separation of peak centers is measured in a frequency domain and quantitative J-modulation techniques, where the couplings are determined by variations in peak intensities induced by both J-couplings and RDC's [55,65]. The frequency displacement techniques can be further categorized into three distinct kinds of experiments.

Firstly in experiments for the direct measurement of the resolved J-couplings from heteronuclear edited spectra for small and medium sized molecules, secondly in spin-state-selection techniques such as the IPAP experiment [66], S³E [67,68] and S³CT [69] scheme for resolved crowded spectra and thirdly in exclusive correlation spectroscopy (E.COSY) type experiments [70].

The latter ones are suitable to resolve small couplings between a pair of spins through a larger coupling between another pair of spins. Moreover, transverse relaxation optimized spectroscopy (TROSY) experiments are widely used in RDC measurements for the studying of large biomolecules or complexes [71]. Some useful pulse sequences for the measurement of RDC's in small and medium sized proteins, including the IPAP [^{1}H-^{15}N]- HSQC, were summarized by Prestegard and coworkers in 2004 [55].

2.5.2.1 IPAP - HSQC - experiment

The in-phase anti-phase (IPAP) HSQC experiment is used to measure RDC between amide ^{1}H and ^{15}N spins and is carried out by acquiring two data sets in an interleaved manner. A representation of the pulse sequence of an IPAP [^{1}H-^{15}N]- HSQC experiment is shown in Fig 8A. In contrast to a standard HSQC experiment, the pulse scheme of an IPAP HSQC optionally includes a spin-echo element (represented as white rectangle in Figure 1.12, left panel) on ^{15}N preceding the t_1 evolution period. In the absence of the spin-echo element, the experiment behaves like a regular HSQC with ^{1}H coupled to ^{15}N during t_1 evolution. The ^{15}N chemical shifts are modulated by cos ($J_{HN}t_1$) and the resulting doublets are in-phase. In the presence of the spin-echo element, the ^{15}N anti-phase term is transferred to an in-phase term prior to the ^{15}N evolution. After the ^{15}N evolution period, only the term modulated by sin ($J_{HN}t_1$) is detected and the generating doublets are in anti-phase. The 90° purge pulse on ^{1}H prior to the t_1 evolution period is used to transfer the remaining residual cos ($\pi J_{HN}\Delta$)H_zN_x to unobservable multiple quantum terms and the two low power 90°$_{-x}$ shape pulses (see Figure 1.12. left panel) surrounding the final ^{1}H 180° pulse are a part of the water signal suppression. A simple pictorial way to visualize the separation of the multiplets is the addition or subtraction of the in-phase and anti-phase spectra (see schematically representation in Figure 1.12, right panel). The resulting two spectra look like regular HSQC's with ^{15}N chemical shifts displaced by either +J_{NH}/2 or -J_{NH}/2. The couplings can then easily be determined by the frequency differences of corresponding peaks. Therefore the separation step of the multiplet components is an interesting way to enhance the spectral resolution and to enable the measurement of small couplings, that are hampered by overcrowding problems [28,55,66,72].
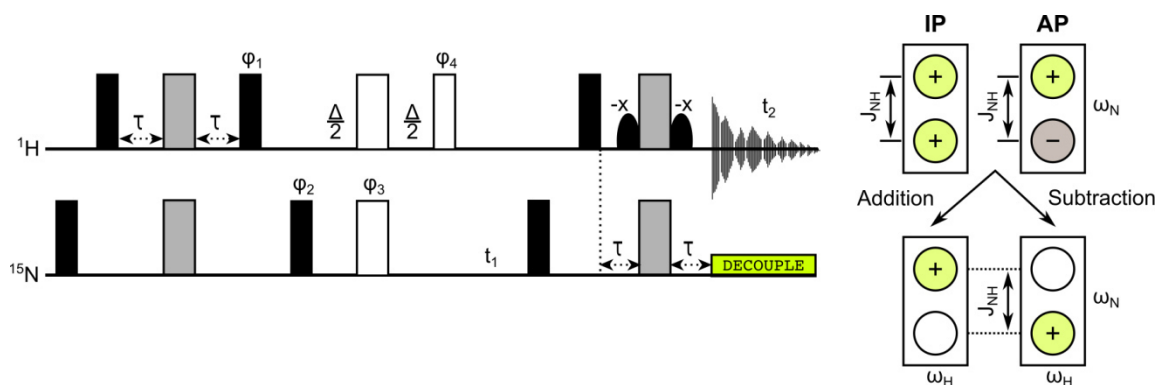
Figure 1.12: IPAP [^{15}N, ^{1}H]- HSQC experiment - Representation of the pulse scheme (left panel). Filled black rectangles and filled grey rectangles represent 90° and 180° flip angles, respectively with phase x, unless indicated. The Δ/2 - 180° (^{1}H /^{15}N) - Δ/2 90°$_{\sigma4}$ sequence, represented by filled white rectangles, will be used in the IPAP [^{15}N, ^{1}H]- HSQC experiment for generating the anti-phase (AP) spectrum and is absent for generating the in-phase (IP) spectrum. The in-phase- and the anti-phase spectra are recorded in an interleaved manner. The two low power 90°$_{-x}$ shape pulses surrounding the final ^{1}H 180° pulse are for water-solvent-suppression. The delay durations are 2τ = 1/(2J$_{NH}$) and Δ = 1/(2J$_{NH}$). During the time t$_{1}$ the decoupling of the ^{13}C nuclei takes place (not shown in the figure). Phase cycling is given by: Φ$_{1}$ = -y, y; Φ$_{2}$ = 2(x), 2(-x) for the in-phase experiment and by Φ$_{1}$ = -y, y; Φ$_{2}$ = 2(y), 2(-y); Φ$_{3}$ = 4(x), 4(y), 4(-x), 4(-y); Φ$_{4}$ = 8(x), 8(-x) for the anti-phase experiment. Receiver = x, 2(-x), x, -x, 2(x), -x for anti-phase [adapted from [66]]. **Schematic representation of a procedure for extracting ^{1}H-^{15}N couplings from an IPAP [^{15}N, ^{1}H]- HSQC experiment by addition or subtraction of the in-phase (IP) and anti-phase (AP) spectra (right panel) [adapted from [72]]**

2.5.3 Alignment Media

As outlined previously, in an isotropic solution, nuclear dipolar couplings are averaged to exactly zero by Brownian rotational diffusion. Therefore the partial alignment of dissolved molecules is clearly essential for the measurement of residual dipolar couplings by solution state NMR. In general, two mechanisms of partial alignment are described in literature. In the first one, high magnetic fields and inherent anisotropies in magnetic susceptibilities of the molecules of interest will be used to directly induce non isotropic distributions whereas in the second case liquid crystalline media will be added to the molecule under study to indirectly induce non isotropic distributions through collisional interactions [55]. In recent years, several types of media have been published to align macro molecules in solution [55,56,65]. The most commonly used media, their properties as well as their advantages and drawbacks are summarized in Table 1.2

2.5.3.1 Bicelles

Disc-shaped bicelles were the first medium used to collect RDC's in biomolecules and were originally developed in 1990 by Prestegard, Sanders and coworkers [73-75]. The most common medium consists of a mixture of phospholipid typical of those found in biological membranes, dimyristoylphosphatidylcholine (DMPC), and a detergent-like lipid, dihexanoylphosphatidyl-choline (DHPC) in low concentration to form planar bicelles (depicted in Figure 1.11 - right panel) where DMPC constitutes the planar bilayer region and DHPC stabilizes the rim of the bicelles [76,77]. The mixture out of DMPC / DHPC was found to be particularly robust and is usually stable in a temperature range of 25 - 45°C [61]. However, temperature studies have shown that bicelles align in the presence of an external magnetic field at 35°C, whereas they are isotropic at 25°C [78].

This can particularly be observed with diluted samples (< 70 mg/ml), which tend to be unstable close to the lower temperature range (25 - 30°C), which then often leads to a separation into an aligned and an isotropic phase [61].

In order to avoid phase separation, small amounts of charged amphiphile such as sodium dodecyl sulfate (SDS) or cetyltrimetylammonium bromide (CTAB) can be added to DMPC/DHPC bicelle preparations. Alternatively, charged phospholipids such as DMP-serine, DMTAP, DMPG or DMPC can be used [61,79]. Charged amphiphiles prevent the aggregation of the phospholipid bilayers by the introduced electrostatic repulsion which also results in an extended temperature range of the stable liquid crystalline phase. Moreover, due to the charge introduction the biomolecule studied will additionally be oriented by electrostatic interactions. Therefore, the partial alignment of dissolved molecules in a bicelle solution can be very complex and includes steric interactions, electrostatic interactions, as well as specific surface associations [76].

However, the formation of a homogenous liquid crystalline phase is not only dependent on temperature, but it also relies on a number of other factors, such as the absolute concentration, the molar ratio of the mixture and the ionic strength [61]. The molar ratio of DMPC/DHPC is a critical factor, notably in solutions with a low bicelle concentration. If the molar ratio is too low, the lipids form small size discs that are too small to generate measurable alignment, whereas a ratio too high can result in non-planar bicellles that may collaps to spherical micelles through DMPC [61]. The preparation of bicelle samples is straightforward and is described in the literature [61,80]. Accordingly, a molar ratio of the lipids of 3 - 3.5 is recommended to form bicelles with a diameter of 200 - 250 Å and a thickness of approximately 40 Å.

2.5.3.2 Filamentous Phages

Filamentous phages are another liquid crystalline medium to align biomolecules and were introduced in parallel in 1998 by Clore *et. al.* and Pardy and coworkers [81,82]. The most widely used phages are the bacteriophage Pf1 and the bacteriophage fd. The tobacco mosaic virus and the phage M13 are other examples of rod shaped viruses that are suitable for the partial alignment of biomolecules [81].

Filamentous phages are highly negatively charged particles. Therefore, the orientation of biomolecules studied is largely induced by electrostatic interactions, making bacteriophages especially useful for the study of negative proteins as well as nucleic acids, including DNA and various types of RNA [55]. Moreover, the strong charge - charge repulsion produces order with minimal line broadening [76].

In comparison to bicelles, the liquid crystalline phase formed by filamentous phages is stable over a wider temperature range (5 - 50°C). Furthermore, the handling of phage-based liquid crystal media is more convenient and the measured protein samples can be easily recovered by high-speed centrifugation of the virus particles.

Limitations with the alignment of biomolecules by phages arises in the presence of high salt concentrations and at low pH values [83]. Recommended pH values for phage solutions are between pH 6.5 and 8. Below pH 6, the partial protonation of aspartate and glutamate sidechains of the phage coat protein reduces the net charge of the virus particles, leading to aggregation and consequently the sample loses liquid crystallinity. In addition high salt concentrations should be avoided at lower phage concentrations ($c_{Pf1} < 13$mg/ml), because they influence the nematic phase boundary, resulting in isotropic samples [61].

Table 1.2: Commonly used alignment media for the measurement of RDC's in biomolecules

media type	composition	charge	temp (°C)	advantages & disadvantages	ref
Liquid crystals					
ester linked phospholipid-bicelles	DMPC/DHPC	neutral	25 - 45	(+) easy preparation (-) degradation at low pH values, expensive (-) not stable in presence of detergents and lipids	77,84
ether linked phospholipid bicelles	DIODPC / CHAPSO	neutral	10 - 55	(+) stable at low pH values (-) not stable at neutral pH values, and in presence of detergents and lipids (-) expensive	85
lamellar liquid crystalline phases	C_mE_n / *n*-hexanol	neutral	0-40	(+) inexpensive, insensitive to pH, low binding affinity for biomolecules	86
purple membrane fragments	2D crystalline fragments of lipids and bacteriorhodopsin	negative	0 - 70	(+) reduced line broadening	87,88
cetylpiridinium - based media / Helfrich phases	CPCl, *n*-hexanol, NaCl CPBr, *n*-hexanol, NaBr	positive	15 - 60	(+) CPCl suitable for high ionic strength (+) CPBr suitable for low ionic strength (-) CPBr is a surfactant	61,89,90
rod-shaped viruses	bacteriophage Pf1 bacteriophage fd tobacco mosaic virus	negative	5 - 60	(+) sample is recoverable (-) loss of liquid crystallinity at pH < 6 and at high ionic strength	81,82,90,91
hydrogels					
stretched / compressed gels	polyacrylamide	neutral	5 - 45	(+) easy sample recovery (+) high compatibility with biomolecules (-) inhomogeneous, line broadening	88,92-94
charged gels	polyacrylamide / acrylate	negative	5 - 45	(+) decreased line broadening (-) complicate preparation	95
Others					
composite media	Combination of bacteriophages and polyacrylamide	negative	-	(+) steric and electrostatic interactions	96
DNA based media	DNA nanotubes G-tetrad DNA	negative	-	(+) compatible against detergents (+) G-tetrad DNA is less expensive (+) partially compatible in presence of detergents (-) expensive	97,98

Chapter II

3. Material and Methods

3.1 Material

Chemicals were purchased by Carl Roth GmbH & Co KG (Karlsruhe, Germany), SERVA Electrophoresis GmbH (Heidelberg, Germany), Sigma-Aldrich Chemie GmbH (Taufkirchen, Germany), VWR International GmbH (Ismaning, Germany) and Roche Diagnostics GmbH (Unterhaching, Germany), unless stated otherwise. Isotopes such as $^{15}NH_4Cl$, ^{13}C-glucose and D_2O were ordered by Euroiso-top (Saint-Aubin, France), Cambridge Isotope Laboratories, Inc (Tewksbury, USA) or Sigma-Aldrich Chemie GmbH (Taufkirchen, Germany).

3.2 Methods

3.2.1 Plasmid preparation

Expression plasmids (pET15b) for *mycobacterium tuberculosis* PknG1-750 (wild type, NORS-RD-KD-TPRD, Uniprot accession no. P9WI73), PknG 74-750 (n-terminal truncated version, RD-KD-TPRD) as well as for PknG 1-147 (NORS-RD) were kindly provided by Dr. Nicole Scherr and Prof. Dr. Jean Pieters from the Biozentrum of the University of Basel. Wild type and PknG mutant constructs had been prepared by inserting the respective coding sequence into a pET15b vector (Novagen) using the Nde1 and XhoI restriction enzyme sites [23]. The use of this plasmid allows to express PknG proteins with an N-terminal six residues long histidine-tag that is embedded into several other residues and followed by a thrombin cleavage site as well as a histidine (MGSSHHHHHHSSGLVPRGSH-).

PknG1-75, PknG74-147, PknG74-420 as well as the N-terminal quadrupole mutant PknG1-147 4C/4S (Cys[106], Cys[109], Cys[128], Cys[131]) were constructed by using the QuickChange site-directed mutagenesis method (Stratagene, La Jolla, CA). Table 2.1 summarizes the used template DNA's as well as the corresponding forward and reverse primers.

The expression plasmid for PknG1-75 was obtained by introducing a stop codon at amino acid position G76, which corresponds to a change of the DNA sequence from 5'-GGC-3' to 5'-TGA-3'. For generating PknG74-420 a stop codon at amino acid position T421 was inserted, corresponding to a change of DNA sequence from 5'-ACA-3' to 5'-TAA-3'.

For the preparation of the PknG 74-147 expression plasmid a factor Xa cleavage site (IEGR) was introduced by mutating residues 70 to 73 (PVRR), corresponding to a change of the DNA sequence from 5'-CCGGTCAGA-3' to 5'-ATCGAAGGC-3'.

In addition, the region containing the thrombin protease site (LVPRGS) was mutated and the following histidine to a factor Xa recognition site (PRGSH to IEGR), corresponding to a change of the DNA sequence from 5'-CCGCGCGGCAGCCAT-3' to 5'-ATCGAAGGCCGC-3'. The construct prepared following that procedure, allows to purify PknG 1-147 with the native N-terminus. For the preparation of the quadrupole mutant, the cystein's of the two CXXCG motifs (motif 1: ^{106}CWNCG^{110} & motif 2: ^{128}CPYCG^{132}) within the rubredoxin domain were mutated to serine's, corresponding to a change of the DNA sequence for motif 1 from 5'-TGCTGGAACTGT-3' to 5'-TCTTGGAACTCT-3' and for motif 2 from 5'-TGTCCCTATTGC-3' to 5'-TCTCCCTATTCT-3'.

Table 2.1: Template DNA and primer for site directed mutagenesis

PknG1-75	
Template DNA	pET15b::PknG1-147
forward primer	5'-GACGGCTGGGC**TGA**GGCCTGGTGGAAATC-3'
reverse primer	5'-GATTTCCACCAGGCC**TCA**GCCCAGCCGTC-3'
PknG74-147	
Template DNA	pET15b::PknG1-147
forward primer	5'-CCAGCCGGGTGCGCCCG**ATCGAAGGC**CGGCTGGGCGGCGGC-3'
reverse primer	5'-GCCGCCGCCCAGCCG**GCCTTCGAT**CGGGCGCACCCGGCTGG-3'
PknG74-420	
Template DNA	pET15b::PknG74-750
forward primer	5'-CCCAGTCGGTCG**TAA**TTTGGAGTGGACCTGCTGG-3'
reverse primer	5'-CCAGCAGGTCCACTCCAAA**TTA**CGACCGACTGGG-3'
PknG1-147 4C/4S	
Template DNA	pET15b::PknG1-147
Forward primer (C^{106},C^{109})	5'- AAGCGGTTC**TCT**TGGAAC**TCT**GGACGTCCCGTC-3'
Reverse primer (C^{106},C^{109})	5'- GACGGGACGTCC**AGA**GTTCCA**AGA**GAACCGCTT-3'
Forward primer (C^{128},C^{131})	5'-GGGAGCTTCAGAGGGCTGG**TCT**CCCTAT**TCT**GGCAGCCCG-3'
Reverse primer (C^{128},C^{131})	5'-CGGGCTGCC**AGA**ATAGGG**AGA**CCAGCCCTCTGAAGCTCCC-3'
PknG1-147 (thrombin to factor Xa cleavage site)	
Template DNA	pET15b::PknG1-147
Forward primer	5'-GCGGCCTGGTG**ATCGAAGGCCGC**ATGGCCAAAGCG-3'
Reverse primer	5'-CGCTTTGGCCAT**GCGGCCTTCGAT**CACCAGGCCGC-3'

The Phusion® High-Fidelity DNA Polymerase used and the restriction enzyme DpnI were obtained from New England Biolabs. The used sample compositions and the temperature profile for site directed mutagenesis are listed in Table 2.2 and Table 2.3. The isolation of plasmid DNA was carried out by using Wizard® Plus SV Miniprep Kit (Promega). The authenticity of the generated mutants was confirmed by DNA sequencing.

Table 2.2: Temperature profile

Segment	Cycles	Temperature	Time
1	1	95°C	1min
2	18	95°C	50s
		60°C	50s
		68°C	6 min*
3	1	68°C	7 min
		4°C	∞

* based on the length of the template vector

Table 2.3: Composition of PCR approaches

Component	Volume [50 µl]
dsDNA template (5-50ng)	0.5 -1 µl
Oligonucleotide primer forward (125 ng)	2.5 µl
Oligonucleotide primer reverse (125 ng)	2.5 µl
50 x dNTP mix (10 mM each)	1 µl
5 x reaction buffer (HF or GC)	10 µl
ddH$_2$O	30,5 - 33 µl
DMSO	0 - 1.5 µl
Pfusion High-Fidelity DNA polymerase	0.5 - 1 µl

3.2.2 Protein Expression

All PknG constructs were finally expressed by *Escherichia coli* BL21 (DE3) cells (Novagen). 100 ml of LB medium supplemented with 100 µg/ml ampicillin were inoculated with a single colony of transformed cells and grown at 37 °C overnight for a large scale culture in rich medium (LB) or throughout the day for a large scale culture in minimal medium (M9). Cultures were usually grown in a shaking flask incubator at 125 rpm. For expression in LB medium supplemented with 100 µg/ml ampicillin the preculture was diluted 1:40. For expression in M9 medium, the preculture was diluted 1:200 in 100 ml unlabeled M9 minimal medium, supplemented with 100 µg/ml ampicillin. The prepared second starter culture was grown at 37°C over night and the next morning diluted in labeled M9 minimal medium so that the starting OD_{600} was about 0.05. The cells were grown at 37°C to an OD_{600} of about 0.4-0.6 and then induced with IPTG as described in the following:

His-PknG 1-147: The protein expression of His-tagged PknG 1-147 in the metal free state was induced with 0.1 mM IPTG for 16 h at 15°C. For the expression of the protein in metal bound state 0.1 mM $ZnCl_2$ were added at the same time as the IPTG. **His-PknG1-75:** Protein expression of His-tagged PknG 1-75 was induced with 1 mM IPTG for 3 h at 37 °C. **His-PknG 74-147:** Protein expression of His-tagged PknG 74-147 was induced with 0.1 mM IPTG for 16 h at 15°C. 0.1 mM $ZnCl_2$ were added at the same time as the IPTG. **His-PknG1-147 4C / 4S:** Protein expression of His-tagged PknG 1-147 4C / 4S was induced with 1 mM IPTG for 2 h at 37 °C. **His-PknG74-420:** Protein expression of labeled His-tagged PknG74-420 was induced with 0.1 mM IPTG for 16 h at 15 °C in M9 media, containing 70 % D_2O.
For *in vitro* kinase assays His-tagged **PknG1-750,** His-tagged **PknG74-750** and His-tagged PknG74-420 were expressed in rich LB medium after the induction with 0.1mM IPTG for 16h at 15°C.

The preparation of the M9 minimal medium was based on a published recipe [32]. The M9 minimal medium used contained M9 salts w/o NH_4Cl, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 1x BME vitamin solution (Sigma), 4 g/l unlabeled glucose or 2 g/l ^{13}C-glucose, and 1 g/l $^{15}NH_4Cl$. Additionally to this recipe, the medium was supplemented with 1x trace element solution (134 µM EDTA, 31 µM $FeCl_3$ x $6H_2O$, 6.2 µM $ZnCl_2$, 0.76 µM $CuCl_2$ x $2H_2O$, 0.42 µM $CoCl_2$ x $2H_2O$, 1.62 µM H_3BO_3, 0.081 µM $MnCl_2$ x $4H_2O$). ^{15}N labeled proteins were prepared in minimal medium containing $^{15}NH_4Cl$ as the sole nitrogen source. Additional ^{13}C labeling was achieved by providing ^{13}C-glucose as the sole carbon source.

3.2.3 Protein extraction and purification

Cells expressing a PknG fragment were harvested by centrifugation at 4 °C (30 min, 4000xg). The cell pellet was resuspended in 50 ml lysis buffer (20 mM Tris, 500 mM NaCl, 20 mM imidazole, 10 mM benzamidine, 50x Roche protease inhibitor cocktail EDTA free, pH 7.5). PMSF (1 mM) was added just before the cell disruption by sonication (Sonopuls Bandelin, UV3200) for 15 min on ice, with a power level of 40 % and a pulse length of 5s. After the sonication the lysate was cleared by a centrifugation for 30 min at 28k rpm and 4°C using an ultracentrifuge. The supernatant was loaded on a gravity flow column filled with 5 ml Ni-NTA agarose beads (Qiagen). The column was washed with ~50 ml buffer containing 20 mM Tris, 500 mM NaCl, 50 mM imidazole, pH 7.5. The His-tagged protein was eluted by increasing the imidazole concentration in the buffer stepwise to 100 mM, 200 mM and finally 500 mM imidazole. For each step 20 ml buffer were used. The collected 10 ml fractions were analyzed by SDS-PAGE using 15 % polyacrylamide gels.

His-PknG1-147. Fractions from the Ni-affinity chromatography containing based on the SDS-PAGE analysis significant amounts of PknG 1-147 were pooled and concentrated at 3500xg at 4°C to 5-10 mg/ml using Amicon® Ultra Centrifugal Filter Units (MWCO 10.000, Merck Millipore). The resulting concentrated protein solution was further purified by size exclusion chromatography using a 200 pg Superdex™ HiLoad™ 16/600 column (GE Healthcare) coupled to an ÄKTA Prime FPLC system (GE Healthcare). Before loading the protein, the column had been equilibrated with 20 mM Tris, 150 mM NaCl pH 7.5. The flow rate was 1 ml/min. His-PknG1-147 eluted at about 80 ml. Fractions containing highly pure protein were pooled and concentrated as described above. If a protein expression was induced without addition of $ZnCl_2$, the refolding of the RD domain in the NMR tube could be achieved by adding 10 mM TCEP and 0.2 mM $ZnCl_2$ if the protein concentration was < 0.2 mM (1 Zn^{2+} per protein molecule needed, data not shown). The protein concentration was determined by UV spectroscopy and the correct molecular weight was confirmed by mass spectrometry.

PknG1-147. The preparation of PknG1-147 with its native N-terminus was only tried from LB medium. Fractions from the Ni-affinity chromatography run containing His-PknG1-147 with a factor Xa site preceding residue 1 were pooled, diluted with 20 mM Tris, 150 mM NaCl, pH 7.5 to a concentration level of 50 mM imidazole and then concentrated at 3500xg at 4°C to 5-10 mg/ml in Amicon® Ultra Centrifugal Filter Units (MWCO 10.000, Merck Millipore).

To further reduce the imidazole and salt concentration, the sample was washed with 20 mM Tris 5 times, 150 mM NaCl pH 7.5 by cycles of (using a dilution factor of 1:5) concentrating each time down to about 1 ml and filled up with buffer to a volume of about 5 ml. The final concentration was measured by UV spectroscopy and was about 1 mg/ml. After the addition of 2 mM $CaCl_2$ from a concentrated stock, the protein solution prepared was digested with 1U factor Xa (New England Biolabs) per 50 µg protein sample buffer 25 °C for about 10 minutes. Size exclusion chromatography may be used to remove the cleaved off His-tag PknG1-147. Alternatively, this may be achieved by a second Ni-affinity chromatography run.

His-PknG1-75: Fractions containing PknG 1-75 were pooled, diluted with 20 mM Tris, 150 mM NaCl pH 7.5 to reduce the imidazole concentration to 50 mM and concentrated at 3500xg at 4°C to 5-10 mg/ml in Amicon® Ultra Centrifugal Filter Units (MWCO 3.000, Merck Millipore). The resulting concentrated protein solution was loaded on a semi-preparative C4 column (Jupiter® 5µm C4 300Å, 250 x 10 mm Phenomenex) coupled to a HPLC system (Äkta Purifier, GE Healthcare) and equilibrated at 10 % HPLC buffer B (90 % acetonitrile/ 0.1 % TFA). HPLC buffer A was 0.1 % TFA in water. For the elution of the protein a gradient from 10 % to 90 % buffer B within 70 min (1.14 %/min) was applied with a flow rate of 4 ml/min. The collected protein fractions were analyzed by 15 % SDS-PAGE. PknG1-75 eluted around 35 % B. The lyophilized protein was resuspended in 20 mM Tris, 150 mM NaCl pH 7.5, washed and concentrated at 3.500xg and 4 °C in Amicon® Ultra Centrifugal Filter Units (MWCO 3.000, Merck Millipore). Instead of using RP-HPLC as a second purification step, we also tried size exclusion chromatography, which was performed as described for His-PknG1-147. PknG1-75 eluted at about 110 ml. In this case the fractions containing PknG1-75 were only pooled and concentrated. The protein concentration was determined by UV spectroscopy and the correct molecular weight was confirmed by mass spectrometry.

PknG74-147: Fractions from the Ni-affinity chromatography run containing His-PknG1-147 with a factor Xa site preceding residue 74 were pooled, diluted with 20 mM Tris, 150 mM NaCl, pH 7.5 to a concentration level of 50 mM imidazole and then concentrated at 3500xg at 4°C to 5-10 mg/ml in Amicon® Ultra Centrifugal Filter Units (MWCO 10.000, Merck Millipore). To further reduce the imidazole and salt concentration, the sample was washed with 20 mM Tris, 150 mM NaCl pH 7.5 5 times by cycles of (using a dilution factor of 1:5) concentrating each time down to about 1 ml and filled up with buffer to a volume of about 5 ml. The final concentration was measured by UV spectroscopy and was about 5-10 mg/ml. The protein solution prepared was digested with 1U factor Xa / mg in sample buffer supplemented with 2 mM $CaCl_2$ at 25 °C overnight.

To separate the His-tagged PknG1-73 fragment from the PknG74-147 fragment (Fig. 1B), the digested protein solution was loaded on a semi-preparative C4 column (Jupiter® 5μm C4 300Å, 250 x 10 mm Phenomenex). The RP-HPLC was performed as described for PknG1-75. The collected protein fractions were analyzed by 15 % SDS-PAGE. PknG74-147 eluted around 46% B. The lyophilized protein sample was refolded in a buffer containing 20 mM Tris, 150 mM NaCl, 2 x molar excess of TCEP, 1.5 x molar excess of $ZnCl_2$ pH 7.5. The protein concentration for refolding was usually about 15-20 μM. Refolded PknG74-174 was washed with the same buffer and concentrated using Amicon® Ultra Centrifugal Filter Units (MWCO 3.000, Merck Millipore) at 3.500xg and 4°C. The protein concentration was determined by UV spectroscopy and the correct molecular weight was confirmed by mass spectrometry.

His-PknG1-147 4C/4S: Fractions containing PknG1-147 4C/4S were pooled, diluted with 20 mM Tris, 150 mM NaCl pH 7.5 to reduce imidazole concentration to 50 mM and concentrated at 3500xg at 4°C to 5-10 mg/ml in Amicon® Ultra Centrifugal Filter Units (MWCO 3.000, Merck Millipore). The resulting concentrated protein solution was loaded on a semi-preparative C4 column (Jupiter® 5μm C4 300Å, 250 x 10 mm Phenomenex). The RP-HPLC was performed as described for PknG1-75. The lyophilized protein was resuspended in 20 mM Tris, 150 mM NaCl pH 7.5, washed and concentrated at 3.500xg and 4 °C in Amicon® Ultra Centrifugal Filter Units (MWCO 3.000, Merck Millipore). The protein concentration was determined by UV spectroscopy and the correct molecular weight was confirmed by mass spectrometry.

His-PknG74-420, His-PknG74-750 & His-PknG1-750: Fractions from the Ni-affinity chromatography containing based on the SDS-PAGE analysis significant amounts of PknG were pooled and concentrated at 3500xg at 4°C to 5-10 mg/ml using Amicon® Ultra Centrifugal Filter Units (MWCO 10.000, Merck Millipore). The resulting concentrated protein solution was further purified by size exclusion chromatography using a 200 pg Superdex™ HiLoad™ 16/600 column (GE Healthcare) coupled to an ÄKTA Prime FPLC system (GE Healthcare). Before loading the protein, the column had been equilibrated with 20 mM Tris, 500 mM NaCl pH 7.5. The flow rate was 1 ml/min. Fractions containing highly pure protein were pooled and concentrated as described above. The protein concentration was determined by UV spectroscopy.

The yield of all tested proteins was usually about 10-15 mg per liter rich medium and about 5-10 mg per liter minimal medium.

3.2.4 NMR sample preparation

The protein concentration of the ^{15}N- and ^{15}N-^{13}C-labeled samples of His-PknG 1-75, His-PknG1-147 in the reduced, Zn-bound state, or PknG 74-147 in the reduced Zn-bound or oxidized, metal-free state in 20 mM Tris (pH 7.5), 150 mM NaCl (95 % H_2O, 5% D_2O) ranged from 0.1 - 0.8 mM in 20 mM Tris, 150 mM NaCl, pH 7.5

The samples of PknG74-147 which was refolded from the lyophilized protein obtained after RP-HPLC purification contained furthermore 0.02-0.06 mM TCEP and 0.02-0.05 mM $ZnCl_2$.

The sample of ^{15}N-^{13}C-PknG74-147 (0.2 mM) for measuring residual dipolar couplings contained ~17 mg/ml PF1 phages (ASLA Biotech). The protein concentration of ^{15}N-D-PknG74-420 in 20 mM Tris (pH 7.5), 500 mM NaCl, 10 mM TCEP, 0.5 mM $MgCl_2$, and 0.5 mM ATP was 0.1 mM.

3.2.5 NMR spectroscopy

NMR spectra were acquired at 298 K on Bruker Avance 500, 600, and 900 MHz spectrometers, the 500 and 900 MHz ones equipped with cryogenic probes. The data were processed with NMRPipe [99] and analyzed using NMRView [100]. The backbone assignment for ^{13}C, ^{15}N, and ^{1}H nuclei were based on three dimensional constant time HNCA [46,101], HNCACB [50,102], CCONH-TOCSY [103-108], HNCO [46,101,109] and HNHA [110,111] spectra.

The $^{13}C^{\alpha}$ and $^{1}H^{\alpha}$ secondary shifts were calculated as the difference between the measured chemical shift value and the random coil value for the respective amino acid [112]. $^{3}J_{HNH\alpha}$ coupling constants were obtained from three-dimensional HNHA spectra [111]. Information about the backbone dynamics were derived from ^{15}N relaxation data including T_1 (spin-lattice relaxation), T_2 (spin-spin relaxation), and ^{1}H-(^{15}N)c-NOE. ^{15}N-^{1}H residual dipolar couplings were obtained from the analysis of ^{15}N-^{1}H-IPAP-HSQC data [66]. The maximal $^{1}D_{N-H}$ for PknG74-147 was 15.7 Hz.

3.2.6 In vitro phosphorylation observed by NMR

NMR samples used to monitor substrate phosphorylation based on ^{1}H-^{15}N-HSQC spectra contained 0.1 mM of ^{15}N-His-PknG1-75 or 1-147 (substrate) in the presence of 1 mM ATP and 5 mM MgCl$_2$ in 20 mM Tris (pH 7.5), 150 mM NaCl, 0.05 % NaN$_3$, 5% D$_2$O, and 10 µM of unlabeled, catalytically active His-PknG74-420. Consecutive spectra were acquired at 298K on a Bruker Avance 500 MHz spectrometer equipped with a cryogenic probe. Overnight incubation of the NMR sample at 310K was done using a thermostated waterbath.

3.2.7 Kinase activity assay

In vitro kinase assays were performed as 25 µl reactions in 20 mM Tris (pH 7.5) supplemented with 150 mM NaCl, 10 mM MgCl$_2$, 2 mM MnCl$_2$, and 40 µM [γ-^{32}P] ATP with an activity of 0.5 µCi. For the activity measurements 0.6 µM kinase (His-PknG1-750 or His-PknG74-750 or His-PknG74-420) and a 5-fold molar excess of substrate (His-PknG1-147 or His-PknG1-75) were incubated at 30 °C for 30 minutes. The kinase reaction was stopped by adding 6x SDS PAGE sample buffer and by boiling the sample at 95 °C for 10 minutes. The kinase and / or the substrate with the RD in the oxidized form were obtained by adding a 40-fold molar excess of H$_2$O$_2$ and EDTA and incubation at 4 °C overnight. The next morning a buffer exchange was carried out using a gravity flow Superdex^{TM} G-25 M PD-10 column (GE-Healthcare). Afterwards the protein solution was concentrated using Amicon® ultra centrifugal filter devices (MWCO 10 kDa) at 10k rpm and 4°C. Aliquots taken during the kinase assays were separated by SDS-PAGE using 15% polyacrylamide gels. Phosphorylation of the substrate was detected by applying a phosphor image screen onto the gel using a Typhoon 9200 PhosphorImager. The analysis of the kinase assay data was done with the program ImageQuant (GE Healthcare).

3.2.8 Molecular dynamics simulations

Full atomistic molecular models of PknG74-420 with the RD in the Fe^{2+} and oxidized states were constructed based on the crystal structure of PknG74-750 in complex with the inhibitor AX20017 (PDB ID: 2PZI) [15]. The loops missing in the crystal structure were modeled using ModLoop [113]. The models were solvated in a water box with Na^+/Cl^- ions, mimicking a 100 mM salt concentration.

The molecular systems comprising *ca*. 62500 atoms, were simulated in an *NPT* ensemble at $T = 310$ K and $p = 101.3$ kPa for 250 ns, using a 2 fs integration time step, and treating long-range electrostatic effects using the Particle Mesh Ewald (PME) approach. Simulations were performed using NAMD 2.9 [114] with the CHARMM27 force field [115] and force field parameters for the Fe^{2+}-4Cys center obtained from the literature[116]. The interaction energy between the PknG protein and the ATP-Mg^{2+} complex was computed at Poisson-Boltzmann (PB) continuum electrostatics level using MEAD [117,118]. The dielectric constant ε for the protein and the ATP-Mg^{2+} complex was set to 4, whereas the solvent was modeled using ε=80. A three-step focusing procedure was used to solve the PB equation numerically with grid dimensions of 189, 189, and 381Å, respectively, and a corresponding grid spacing of 1.0Å, 0.4Å, and 0.1Å.Visual Molecular Dynamics (VMD) [119] was used for analyzing the MD trajectories.
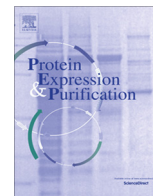
Chapter III

Summary 1st publication

Expression and purification of the natively disordered and redox sensitive metal binding regions of *Mycobacterium tuberculosis* protein kinase G

Matthias Wittwer, Sonja A. Dames

Mycobacterium tuberculosis protein kinase G (PknG) is a eukaryotic-like serine/threonine kinase consisting of four functional domains which is essential for the survival of the pathogen in the human host. The N-terminal regions are responsible for regulating the kinase activity of the catalytic domain (aa 147-405). They consist of the non-regulatory secondary structure (NORS aa 1-75) region harboring the only autophosphorylation residue T63, followed by the redox sensitive rubredoxin domain (RD aa 74-147). As these domains are highly flexible, NMR is the most suitable method for the analysis of their regulatory function. For this purpose, a variety of samples containing individual domains or multidomain constructs are investigated.NMR depends on demanding sample conditions in terms of isotope labeling, quantity and concentration. In this publication optimized preparation conditions of various PknG constructs suitable for NMR investigations are presented. This includes the generation of different expression plasmids as well as the expression and purification of the respective proteins. By applying site directed mutagenesis on the plasmid for His-tagged PknG1-147 the following plasmids were obtained: His-PknG1-75, by introducing a stop codon at position G76, PknG74-147, by introducing a factor Xa recognition site before position 74, and a modified expression plasmid for PknG1-147, by mutating the N-terminal thrombin to a factor Xa recognition site. For all fragments protocols for the expression of isotopically labeled protein in M9 minimal medium, purification to high purity as controlled by SDS PAGE-analysis as well as sufficient yield and concentration for NMR spectroscopy are established. It was shown that the addition of $ZnCl_2$ before induction is necessary for the expression of the RD in its folded, metal bound state. ^{1}H-^{15}N-HSQC spectra of the PknG constructs feature well dispersed backbone signals, which are amenable for sequential assignment. The initial NMR experiments demonstrate the feasibility of structural and dynamic characterization of the regulatory N-terminus, facilitating binding studies of potential drugs against tuberculosis.

Matthias Wittwer designed and conducted experiments in the wet lab, carried out data analysis and was involved in writing the manuscript

Expression and purification of the natively disordered and redox sensitive metal binding regions of *Mycobacterium tuberculosis* protein kinase G

Matthias Wittwer [a], Sonja A. Dames [a,b,∗]

[a] *Technische Universität München, Department of Chemistry, Biomolecular NMR Spectroscopy, Lichtenbergstr. 4, 85747 Garching, Germany*
[b] *Institute of Structural Biology, Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany*

A B S T R A C T

Mycobacterium tuberculosis protein kinase G (PknG) is secreted into host macrophages to block lysosomal degradation. The catalytic domain (∼147–405) is C-terminally flanked by a tetratricopeptide repeat domain (TPRD). The preceding rubredoxin-like metal-binding motif (RD, ∼74–147) mediates PknG redox regulation. The N-terminal ∼75 residues were predicted to show no regulatory secondary structure (NORS) and harbor the only site (T63) phosphorylated *in vivo*. Deletions or mutations in the NORS or the redox-sensitive RD significantly decrease the survival function. Here, we show that the RD appears only to be present in the folded, metal-bound state if $ZnCl_2$ is added upon induction of protein expression in minimal medium. Since factor Xa cleaves at the end of its recognition site (IEGR), a modified expression plasmid for PknG1–147 was obtained by mutating the N-terminal thrombin to a factor Xa recognition site. This allows preparing PknG1–147 with its native N-terminus. We further present a fast approach to generate expression plasmids for only the NORS or the RD by site-directed mutagenesis of the expression plasmid for His-tagged PknG1–147. An expression plasmid for PknG1–75 was obtained by introducing a stop codon at position 76 and one for PknG74–174 by introducing a factor Xa recognition site before position 74. SDS–PAGE analysis shows that all fragments are highly expressed in *E. coli* and can be purified to high purity. Thereby, the established preparation protocols pave the route for the NMR structural characterization of PknG regulation by its N-terminal regions, which is demonstrated by the recorded initial ^1H–^{15}N-HSQC spectra.

© 2015 Elsevier Inc. All rights reserved.

Introduction

Protein kinase G (PknG)[1] is one of eleven eukaryotic-like serine/threonine protein kinases encoded in the genome of *M. tuberculosis*, the causative agent of tuberculosis [1,2]. PknG promotes cellular survival of pathogenic mycobacteria by blocking their lysosomal delivery and thus degradation [3]. Since PknG is secreted into the host macrophages, it is an attractive drug target [3]. PknG is 750 residues long and contains different functional regions (Fig. 1A). The N-terminal about 75 residues were suggested to be intrinsically disordered and thus to show no regulatory secondary structure (NORS). However, this region harbors the only site (T63) that has been shown to be phosphorylated *in vivo* [4]. Mycobacterial survival in host macrophages has further been suggested to require autophosphorylation [5]. Genetic and structural analysis revealed that PknG is the only mycobacterial kinase in which the catalytic domain is N-terminally flanked by a redox-sensitive rubredoxin motif (RD) containing two conserved C–X–X–C-motifs [2,6]. The RD interacts with both, the N-terminal and the C-terminal lobes, of the kinase [6]. Mutations of cysteines in the conserved C–X–X–C motifs to serine or alanine influence the catalytic activity and render PknG insensitive to a regulation by redox changes [4,5]. The C-terminus consists of a tetratricopeptide repeat (TPR) domain, a structural motif typically involved in protein–protein interactions [6,7]. Since deletion of the TPRD reduces the catalytic activity of PknG, it may regulate PknG function by dimerization [4,6].
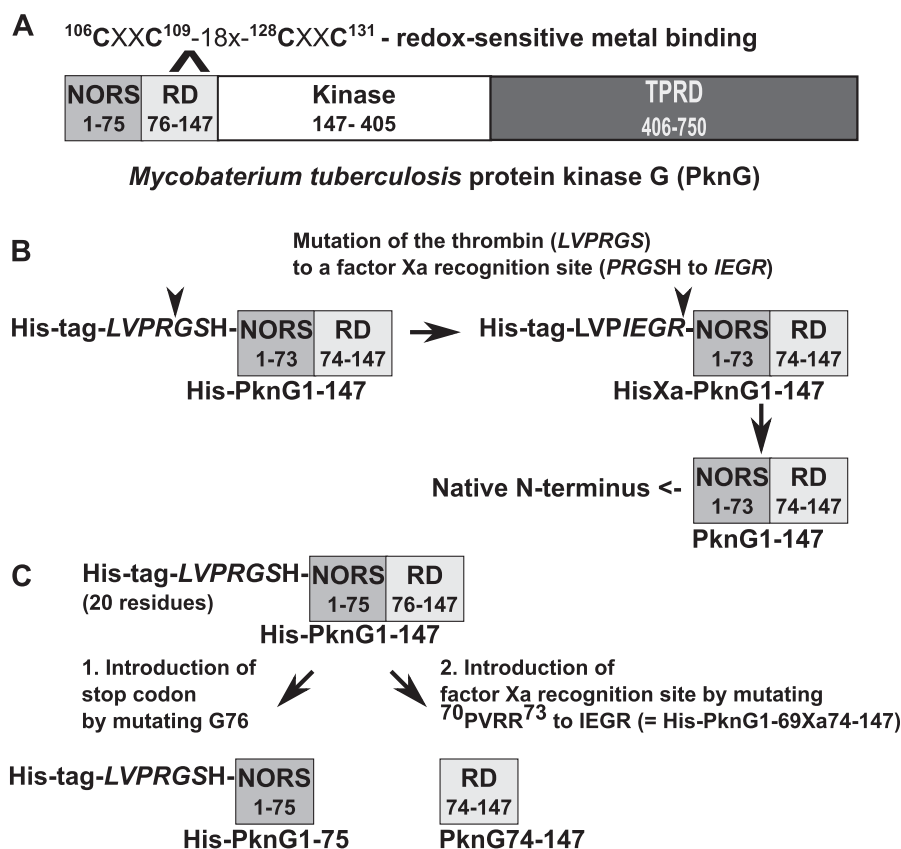
Fig. 1. Domain structure of *mycobacterium tuberculosis* protein kinase G (PknG) and schematic representation of the generation of different expression plasmids for the N-terminal regions based on a site-directed mutagenesis approach. A: PknG contains different functional regions. The N-terminal ~75 residues are predicted to show no regular secondary structure (NORS). The following redox-sensitive rubredoxin-like domain (RD) is N-terminal of the catalytic ser/thr kinase domain that is C-terminally flanked by a tetratricopeptide repeat domain (TPRD). The metal coordinating CXXC motifs of the RD are indicated above. B: To be able to prepare PknG1–147 with the native N-terminus, we prepared a construct in which the N-terminal thrombin site and the subsequent histidine were mutagenized to a factor Xa site. Since factor Xa cleaves at the end of the recognition site, the resulting PknG1–147 fragments starts with the native N-terminal methionine (M1). C: Generation of expression plasmids for only the NORS region or the RD based on the available expression plasmid for His-tagged PknG1–147 (His-PknG1–147) by sited directed mutagenesis. The one for the preparation of His-PknG1–175 was obtained by introducing a stop codon at position 76 and the one for PknG74–174 by replacing residues 70–73 by a factor Xa recognition site.

In order to be able to structurally characterize the N-terminal natively disordered region and the role of the RD for the redox regulation of PknG by multidimensional, heteronuclear NMR techniques, we established expression and purification protocols for different size constructs of these regions. Depending on the expression conditions for PknG1–147 (Fig. 1), ^{1}H–^{15}N-HSQC spectra indicate that the RD is either present in the folded, metal bound form or appears as the NORS region unfolded. To facilitate the NMR resonance assignment and to be able to use the NORS and the RD separately for binding studies, tag-free PknG1–147 and shorter fragments encompassing only residues 1–75 or 74–147, respectively, were obtained by introducing either a factor Xa recognition site or a stop codon at suitable positions in the expression plasmid for His-tagged PknG1–147. The ^{1}H–^{15}N HSQC spectra of PknG1–75 and 74–147 represent almost completely subspectra of the resonances of the respective regions in PknG1–147. All proteins express well in minimal medium, which allows obtaining high yields of isotope labeled, highly pure proteins.

Materials and methods

Plasmid preparation

An expression plasmid for *M. tuberculosis* PknG1–147 (Uniprot accession No. P9WI73) was kindly provided by Dr. Nicole Scherr and Prof. Dr. Jean Pieters from the Biozentrum of the University of Basel. This construct had been prepared by inserting the respective coding sequence into a pET15b vector (Novagen) using the Nde1 and XhoI restriction enzyme sites [5]. Use of this plasmid allows to prepare PknG1–147 with an N-terminal six residues long histidine-tag that is embedded by several other residues and followed by a thrombin site and a histidine (MGSSHHHHHHSSGLV PRGSH–, Fig. 1B). To obtain an expression plasmid that allows to prepare PknG1–147 with its native N-terminus, we mutated the region containing the thrombin protease site (LVPRGS) and the following histidine to a factor Xa recognition site (PRGSH to IEGR → HisXa-PknG1–147, Fig. 1B), corresponding to a change of the DNA sequence from 5′-CCGCGCGGCAGCCAT-3′ to 5′-ATCGA AGGCCGC-3′. The sequences of the used forward and backward PCR primers were 5′-GCGGCCTGGTG**ATCGAAGGCCGC**ATGGCCA AAGCG-3′ and 5′-CGCTTTGGCCAT**GCGGCCTTCGAT**CACCAGGCCGC-3′, respectively. Expression plasmids to prepare His-tagged PknG1–75 (His-PknG1–75) and tag-free PknG74–147 (PknG74–147) were constructed by site directed mutagenesis (Fig. 1C). The one for His-PknG1–75 was obtained by introducing a stop codon at amino acid position G76, corresponding to a change of the DNA sequence from 5′-GGC-3′ to 5′-TGA-3′. The sequences of the used forward and backward PCR primers were 5′-GACGG CTGGGC**TGA**GGCCTGGTGGAAATC-3′ and 5′-GATTTCCACCAGGCC **TCA**GCCCAGCCGTC-3′, respectively. For the preparation of the PknG74–147 expression plasmid a factor Xa cleavage site (IEGR) was introduced by mutating residues 70–73 (PVRR), corresponding to a change of the DNA sequence from 5′-CCGGTCAGA-3′ to 5′-ATCGAAGGC-3′. The sequences of the used forward and backward

PCR primers were 5′-CCAGCCGGGTGCGCCCG**ATCGAAGG**CCGGCT GGGCGGCGGC-3′ and 5′-GCCGCCGCCCAGCCG**GCCTTCGAT**CGGGCG CACCCGGCTGG-3′, respectively. The used Phusion® High-Fidelity DNA Polymerase and the restriction enzyme DpnI were obtained from New England Biolabs. Isolation of plasmid DNA was carried out by using a Wizard® Plus SV Miniprep Kit (Promega). The success of the mutagenesis was verified by DNA sequencing.

Physicochemical parameters of PknG constructs (UniProt accession No. P9WI73)

His-PknG1–147
=PknG1–147 including an N-terminal His-tag and a thrombin recognition site:
MGSSHHHHHHSSGLVPRGSH-PknG1–147, 20 + 147 = 167 residues
Mw = 18,001 Da, pI = 6.22, ε_{280nm} = 13,980 M^{-1} cm^{-1}.
HisXa-PknG1–147
=PknG1–147 including an N-terminal His-tag and a factor Xa recognition site:
MGSSHHHHHHSSGLVIEGR-PknG1–147, 19 + 147 = 166 residues
Mw = 17,923 Da, pI = 5.94, ε_{280nm} = 13,980 M^{-1} cm^{-1}.
His-PknG1–69Xa74–147
=PknG1–147 including an N-terminal His-tag and a thrombin recognition site as well as a factor Xa recognition site replacing residues 70–73:
MGSSHHHHHHSSGLVPRGSH-PknG1–69-IEGR-74–147, 20 + 147 = 167 residues
Mw = 17,949 Da, pI = 5.86, ε_{280nm} = 13,980 M^{-1} cm^{-1}.
PknG1–147
147 residues, Mw = 15,838 Da, pI = 4.99, ε_{280nm} = 13980 M^{-1} cm^{-1}.
His-PknG1–75
=PknG1–75 including an N-terminal His-tag and a protease recognition site:
MGSSHHHHHHSSGLVPRGSH-PknG1–75, 20 + 75 = 95 residues
Mw = 10,241 Da, pI = 7.11, ε_{257nm} = 600 M^{-1} cm^{-1}.
PknG1–75
75 residues, Mw = 8078 Da, pI = 5.67, ε_{257nm} = 600 M^{-1} cm^{-1}.
PknG74–147
74 residues, Mw = 7949 Da, pI = 4.61, ε_{280nm} = 13,980 M^{-1} cm^{-1}.

Protein expression

All PknG constructs were finally expressed from *Escherichia coli* BL21 (DE3) cells (Novagen). For some we tried initially also *Escherichia coli* Rosetta 2 (DE3) (Novagen), however without any gain regarding the expression yield. 100 ml of LB medium supplemented with 100 µg/ml ampicillin (in case of using Rosetta cells additionally supplemented with 40 µg/ml chloramphenicol) were inoculated with a single colony of transformed cells and grown at 37 °C for 8–16 h. Cultures were usually grown in a shaking flask incubator at 125 rpm. For large scale expression in LB medium the overnight starter culture was diluted 1:40. For expression in minimal medium, the LB starter culture that had been growing during the day was diluted 1:200 in 100 ml unlabeled M9 minimal medium and grown at 37 °C over night. The resulting minimal medium starter culture was diluted such in labeled M9 minimal medium that the starting OD_{600} was about 0.05 and grown at 37 °C. All cultures were grown to an OD_{600} of about 0.4–0.6 and induced with IPTG as described in the following:

His-PknG1–147, HisXa-PknG1–147, His-PknG1–69Xa74–147
Expression of the metal free state was induced with 0.1 mM IPTG overnight (~16 h) at 15 °C. For expression of the protein in the metal bound state 0.1 mM $ZnCl_2$ were added at the same time as the IPTG.

His-PknG1–75
Protein expression of His-tagged PknG1–75 was induced with 1 mM IPTG for 3 h at 37 °C.
The used M9 minimal medium contained M9 salts w/o NH_4Cl, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 1× BME vitamin solution (Sigma),

4 g/l unlabeled glucose or 2 g/l ^{13}C-glucose, and 1 g/l $^{15}NH_4Cl$ [8] as well as trace elements (134 µM EDTA, 31 µM $FeCl_3 \cdot 6H_2O$, 6.2 µM $ZnCl_2$, 0.76 µM $CuCl_2 \cdot 2H_2O$, 0.42 µM $CoCl_2 \cdot 2H_2O$, 1.62 µM H_3BO_3, 0.081 µM $MnCl_2 \cdot 4H_2O$, http://www.embl.de/pepcore/pep-core_services/protein_expression/ecoli/n15). ^{15}N labeled proteins were prepared in minimal medium containing $^{15}NH_4Cl$ as the sole nitrogen source. Additional ^{13}C labeling was achieved by providing ^{13}C-glucose as the sole carbon source.

Protein analysis

Protein samples taken along the expression and purification were analyzed by SDS–PAGE using 15% polyacrylamide gels. The used molecular weight markers were Page Ruler Prestained Protein Ladder 10–170 kDa from Thermoscientifc (Figs. 2 and 4) and Roti®-Mark 10–150 from Carl Roth® (Fig. 3). The correct mass of the purified proteins was verified by mass spectrometry and the protein concentration was determined by UV spectroscopy.

Protein extraction and purification

Cells expressing either PknG fragment were harvested by centrifugation at 4 °C (30 min, 4000×g). The cell pellet was resuspended in 50 ml lysis buffer (20 mM Tris, 500 mM NaCl, 20 mM imidazole, 10 mM benzamidine, 50× Roche protease inhibitor cocktail EDTA free, pH 7.5). PMSF (1 mM) was added just before cell disruption by sonication (Sonopuls Bandelin, UV3200) for 15 min on ice with a power level of 40% and a pulse length of 5 s. The lysate was centrifuged for 30 min at 28 k rpm and 4 °C using an ultracentrifuge. The supernatant was loaded on a gravity flow column filled with 5 ml Ni–NTA agarose beads (Qiagen). The column was washed with ~50 ml 20 mM Tris, 500 mM NaCl, 50 mM imidazole, pH 7.5. The His-tagged protein was eluted by stepwise increasing the imidazole concentration to 100 mM, 200 mM, and finally 500 mM imidazole. For each step 20 ml buffer were used. 10 ml fractions were collected. Fractions containing, based on the SDS–PAGE analysis, significant amounts of either PknG protein were pooled and concentrated at 3500×g at 4 °C to 5–10 mg/ml using Amicon® Ultra Centrifugal Filter Units (MWCO 10000 for His-/HisXa-PknG1–147/His-PknG1–69Xa74–147 and MWCO 3000 for PknG1–75 and 74–147, Merck Millipore).

His-PknG1–147
The protein was further purified by size exclusion chromatography using a 200 pg Superdex™ HiLoad™ 16/600 column (GE Healthcare) coupled to an ÄKTA Prime FPLC system (GE Healthcare). Before loading the protein, the column had been equilibrated with NMR buffer (20 mM Tris, 150 mM NaCl, pH 7.5). The flow rate was 1 ml/min. His-PknG1–147 eluted at ~80 ml. Fractions containing highly pure protein were pooled and concentrated as described above. If protein expression was induced without addition of $ZnCl_2$, refolding of the RD domain in the NMR tube could be achieved by adding 10 mM TCEP and a slightly higher than equimolar amount of $ZnCl_2$.

HisXa-PknG → PknG1–147
Preparation of PknG1–147 with its native N-terminus was only tried from LB medium. Fractions from the Ni-affinity chromatography run containing His-PknG1–147 with a factor Xa site preceding residue 1 (HisXa-PknG1–147, Fig. 1B) were pooled, diluted with NMR buffer to a concentration level of 50 mM imidazole and concentrated at as described above. To further reduce the imidazole and salt concentration, the sample was 5 times washed with 20 mM Tris, 150 mM NaCl, pH 7.5 by cycles of (using a dilution factor of 1:5) concentrating each time down to ~1 ml and filling up with buffer to a volume of ~5 ml. The washed and
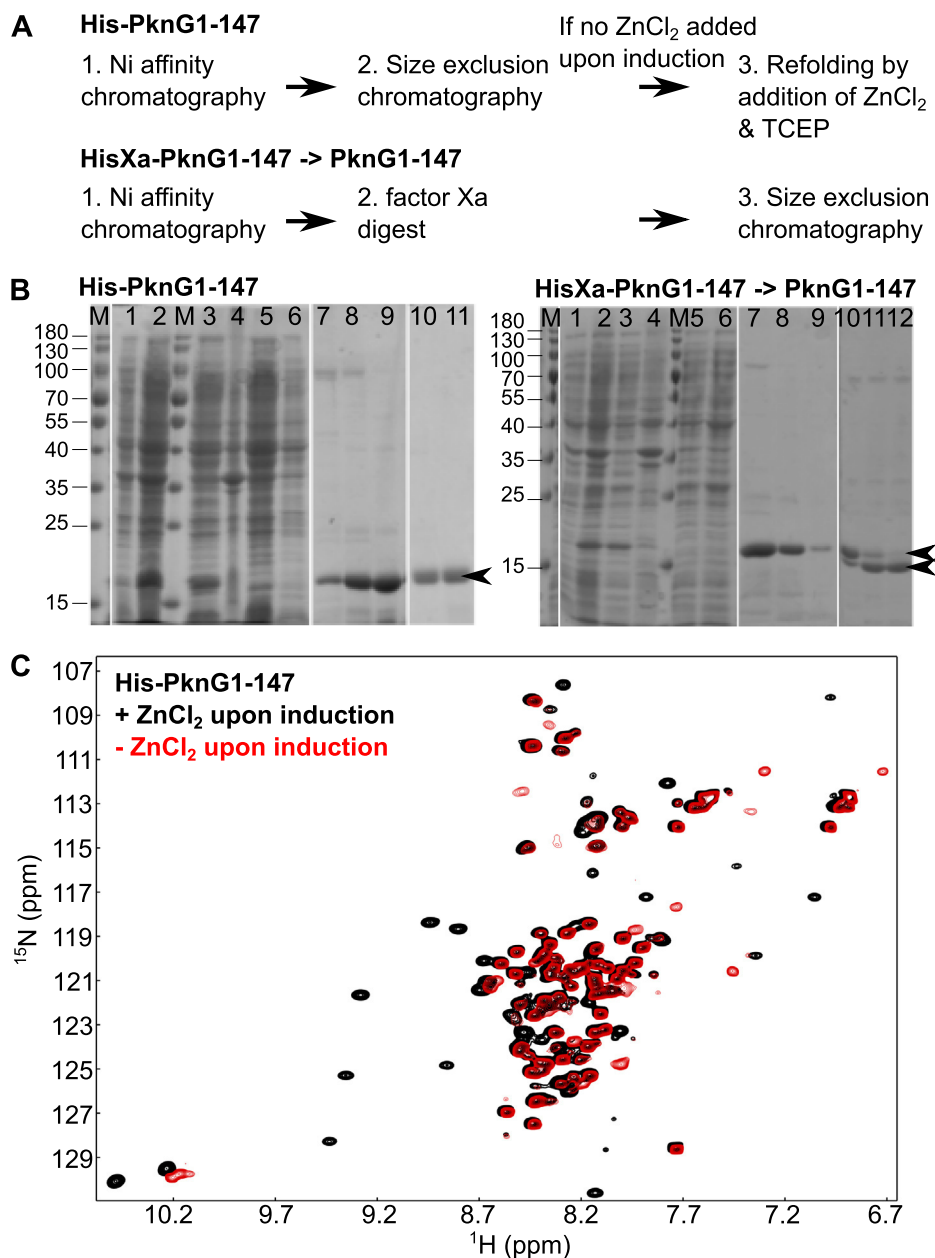
Fig. 2. Expression and purification of His-PknG1–147 and tag-free PknG-1–47 analyzed by SDS–PAGE and NMR analysis of purified isotope labeled His-PknG1–147. A: Flow diagrams of the purification procedures. B: The positions of the SDS–PAGE bands for His-PknG1–147 (left picture) and HisXa-PknG1–147 and PknG1–147 (right picture) are indicated by short black arrows. M, molecular weight marker; Left plot (His-PknG1–147): 1–2, M9 minimal medium culture before and after induction; 3–4, supernatant and pellet after sonication; 5–9, purification of the protein by Ni-affinity chromatography; 5, flow-through after loading; 6, wash fraction; 7–9, representative elution fractions, 10–12, representative fractions of the protein eluted from the size exclusion column; Right plot (HisXa-PknG1–147): 1–9, as left plot but for protein expressed in rich medium; 10–12, factor Xa digest; 10, before; 11–12 after 5 and 10 min. C: Superposition of the ^1H–^{15}N HSQC spectra of ^{15}N-^{13}C-His-PknG1–147 purified from cultures supplemented with 0.1 mM $ZnCl_2$ upon induction (black) and of ^{15}N-labeled protein obtained from cultures only induced with IPTG (red). Indicated by several well-dispersed backbone signals (e.g. ∼8.7–9.7 ppm), the RD regions appears to be present in the folded, metal bound state in the black spectrum. Because the signal dispersion of the backbone amides of the whole red spectrum is rather narrow (∼7.7–8.7 ppm), the RD appears in this sample as unfolded as the preceding natively disordered NORS region. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

concentrated protein solution (∼1 mg/ml) was supplemented with 2 mM $CaCl_2$ and digested with factor Xa (New England Biolabs, 1 U per 50 μg substrate protein) at 25 °C for 10 min. The cleaved off His-tag can be removed by size exclusion chromatography (see His-PknG1–147).

His-PknG1–75

The concentrated protein solution of Ni-affinity chromatography purified protein was loaded on a semi-preparative C4 column (Jupiter® 5 μm C4 300 Å, 250 × 10 mm, Phenomenex) coupled to a HPLC system (Äkta Purifier, GE Healthcare) and

equilibrated at 10% HPLC buffer B (90% acetonitrile/0.1% TFA). HPLC buffer A was 0.1% TFA in water. For the elution of the protein, a gradient from 10% to 90% buffer B was applied with a flow rate of 4 ml/min within 70 min (= 1.14 %/min). His-PknG1–75 eluted at ∼35% B. The protein was lyophilized and resuspended in NMR buffer and washed and concentrated as described above. Instead of RP-HPLC as second purification step, we also tried size exclusion chromatography, which was performed as described for His-PknG1–147. His-PknG1–75 eluted at about 110 ml. In this case the fractions containing His-PknG1–75 were only pooled and concentrated to prepare the NMR sample.
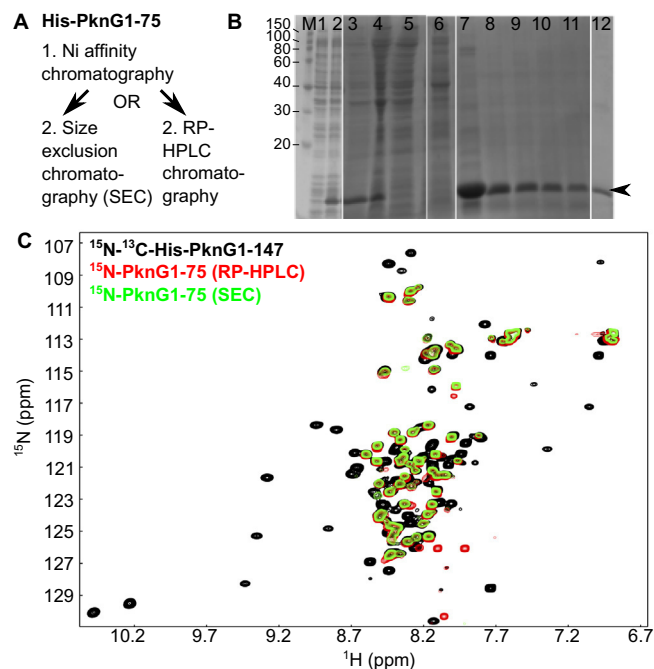
Fig. 3. Expression and purification of His-PknG1–75 analyzed by SDS–PAGE and NMR analysis of the purified isotope labeled protein. A: Flow diagram of the purification procedure. B: The position of the SDS–PAGE band for His-PknG1–75 is indicated by a short black arrow. M, molecular weight marker; 1–2, M9 minimal medium culture before and after induction; 3–4, supernatant and pellet after sonication; 5–11, purification of the protein by Ni-affinity chromatography; 5, flow-through after loading; 6, wash fraction; 7–11, representative elution fractions; 12, RP-HPLC purified protein in NMR buffer. C: Superposition of the ^{1}H–^{15}N HSQC spectra of ^{15}N–^{13}C-His-PknG1–147 purified from cultures supplemented with 0.1 mM ZnCl$_2$ upon induction (black) and of ^{15}N-PknG1–75 purified as second purification step either with RP-HPLC (red) or size exclusion chromatography (green). The red and the green spectrum are largely identical. The signals of both overlap overall very nicely with a subset of peaks of PknG1–147, indicating that the NORS region is similarly unstructured in all 3 samples. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

His-PknG1–69Xa74–147 → PknG74–147

Fractions from the Ni-affinity chromatography run containing His-PknG1–147 with a factor Xa site preceding residue 74 (His-PknG1–69Xa74–147, Fig. 1C) were pooled, concentrated and washed as described for HisXa-PknG1–147. The concentrated protein solution (5–10 mg/ml) was supplemented with 2 mM CaCl$_2$ and digested with factor Xa (1 U per mg substrate protein) at 25 °C overnight. To separate the His-tagged PknG1–69Xa fragment from the PknG74–147 fragment (Fig. 1C), the digested protein solution was loaded on a semi-preparative C4 column (Jupiter® 5 μm C4 300 Å, 250 × 10 mm Phenomenex). The RP-HPLC was performed as described for PknG1–75. PknG74–147 eluted at ~46% B. The fraction containing PknG74–147 was lyophilized and resuspended in NMR buffer and refolded by adding a 2-fold molar excess of TCEP and a 1.5-fold molar excess of ZnCl$_2$. The protein concentration for refolding was usually about 15–20 μM. Refolded PknG74–174 was several times washed with the used refolding buffer (NMR buffer + 22.5–30 μM ZnCl$_2$ + 30–40 μM TCEP) and concentrated as described above.

The final protein yield was usually about 10–15 mg from 1 liter rich medium and about 5–10 mg from 1 liter minimal medium.

NMR spectroscopy

NMR samples contained ~100–800 μM protein in NMR buffer supplemented with 5% D$_2$O and where applicable other supplements (TCEP & ZnCl$_2$ – see refolding of PknG1–147/74–147).
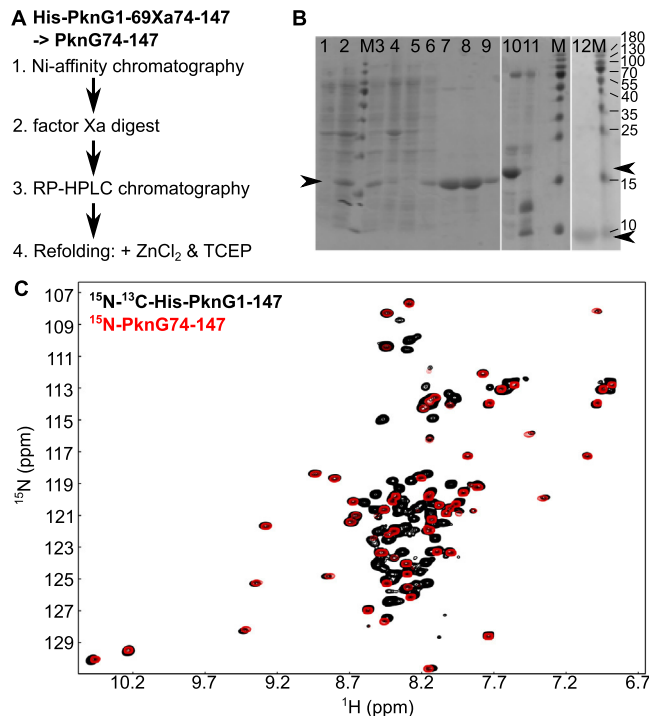

Fig. 4. Expression and purification of PknG74–147 analyzed by SDS–PAGE and NMR analysis of the purified isotope labeled protein. A: Flow diagram of the purification procedure. B: The positions of the SDS–PAGE bands for His-PknG1–69Xa74–147 and PknG74–147 are indicated by short black arrows. M, molecular weight marker; 1–2, M9 minimal medium culture before and after induction; 3–4, supernatant and pellet after sonication; 5–9, purification of the protein by Ni-affinity chromatography; 5, flow-through after loading; 6, wash fraction; 7–9, representative elution fractions, 10–12, representative elution fractions from the size exclusion chromatography run; 10–11, factor Xa digest; 10, before; 11 after ~16 h; 12, representative fraction of the RP-HPLC purified protein. C: Superposition of the ^{1}H–^{15}N HSQC spectra of ^{15}N–^{13}C-His-PknG1–147 purified from cultures supplemented with 0.1 mM ZnCl$_2$ upon induction (black) and of RP-HPLC purified, refolded ^{15}N-PknG74–147 (red). Indicated by the almost perfect overlap of several well-dispersed backbone signals (e.g. ~8.7–9.7 ppm), the RD regions appears to adopt the same folded, metal bound state in both proteins. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

NMR spectra were acquired at 298 K on a Bruker Avance 500 spectrometer equipped with a cryogenic probe. Data were processed with NMRPipe [9] and analyzed using NMRView [10].

Results and discussion

Expression, purification, and NMR analysis of unfolded and metal bound His-PknG1–147

His-PknG1–147 (Fig. 1B) was highly expressed from *E. coli* BL21 (DE3) cells (Fig. 2B left plot, lanes 1–2). The protein was also well expressed if *E. coli* Rosetta (DE3) cells were used (data not shown). Since the use of Rosetta cells did not increase the yield, however affords the use of a second antibiotic in the medium, we used for all subsequent purifications BL21 (DE3) cells. After expression at 15 °C over night, the protein was mostly found in the soluble fraction (Fig. 2B left plot, lanes 3–4). The presence of the natively disordered N-terminal NORS region (Fig. 1B) did not result in strong bands for potential degradation products in the soluble fraction (Fig. 2B left plot, lane 3) or in the elution fractions of the subsequent purification steps (Fig. 2B left plot, lanes 7–11). Extraction of His-PknG1–147 from the centrifuged cell lysate by Ni-affinity chromatography (Fig. 2B left plot, lanes 5–9) resulted already in

rather clean protein in the elution fractions (Fig. 2B left plot, lanes 7–9). During the subsequent size exclusion chromatography run the remaining few impurities could be almost completely removed (Fig. 2B left plot, lanes 10–11).

Fig. 2C shows in red the $^1H–^{15}N$-HSQC spectrum of ^{15}N-His-PknG1–147 that had been purified from M9 minimal medium cultures that had only been induced with IPTG. The spectrum that is characterized by low signal dispersion is overall typical for an unfolded protein. Since the RD domain may be able to fold to the metal-bound rubredoxin fold observed in the crystal structure of PknG74–750 [6] in the absence of the kinase domain, we tested if we could obtain His-PknG1–147 with the RD in the folded, metal bound state by adding a divalent metal ion upon induction [11]. Based on the black $^1H–^{15}N$-HSQC spectrum in Fig. 2C, adding 0.1 mM $ZnCl_2$ upon induction of protein expression with IPTG was sufficient to obtain His-PknG1–147 with the RD in folded state. This is indicated by the appearance of several well-dispersed backbone amide crosspeaks, e.g. in the range from 8.7 to 9.7 ppm, as well as for example significantly different peak positions for the side chain amide protons of the tryptophan indole rings (W107 & W127) around 10–10.5 ppm. The same spectral appearance can be obtained if a reducing agent (e.g. 10 mM TCEP) and $\geqslant 1$ equivalent of Zn^{2+} per protein molecule (e.g. by adding $ZnCl_2$ from a 50–100 mM stock solution) are added to the protein purified from cells that were induced without adding $ZnCl_2$ (data not shown).

Generation of a construct for the preparation of PknG1–147 with its native N-terminus and expression and purification of the respective protein

Because the presence of the native N-terminus is sometimes vital for the protein function, for example that of cadherins mediating cell adhesion [12], and because removal of the His-tag may facilitate the NMR resonance assignment and/or be beneficial for future interactions studies, we prepared further an expression plasmid that allows to prepare PknG1–147 with its native N-terminus (PknG1–147). For this, we replaced the N-terminal thrombin site that is followed a histidine by a factor Xa site (HisXa-PknG1–147, Fig. 1B) using site-directed mutagenesis. Compared to thrombin that cleaves between the R and the G of its recognition sequence (LVPR^GS) thereby leaving two non-native residues (GS), factor Xa cleaves at the end of its recognition sequence (IEGR^) thereby leaving no non-native residues. Also this PknG protein expressed well in *E. coli* cells (Fig. 2B right plot, lanes 1–2) and was almost completely detected in the soluble fraction (Fig. 2B right plot, lanes 3–4). The already rather clean protein obtained from the Ni-affinity purification (Fig. 2B right plot, lanes 5–9) was digested with factor Xa (Fig. 2B right plot, lanes 10–12) to result in complete removal of the His-tag.

Generation of a construct for the preparation of NORS region and expression and purification of the resulting His-PknG1–75 protein

Since cloning of a new expression plasmid from scratch can be rather time consuming, we thought of a faster approach to obtain an expression plasmid for only the N-terminal NORS region (Fig. 1A). Because site-directed mutagenesis involving only few DNA sequence changes is generally highly efficient, we just incorporated a stop codon at the position of residue 76 in the expression plasmid for His-PknG1–147 (Fig. 1C). As His-PknG1–147 (Fig. 1C), the construct including a stop codon after residue 75 corresponding to His-PknG1–75 was also highly expressed from *E. coli* BL21 (DE3) cells (Fig. 3B, lanes 1–2). After expression at 37 °C for 3 h, the majority of the protein was found in the soluble fraction (Fig. 3B, lanes 1–4). Based on the SDS–PAGE analysis, the expression and purification of the His-tagged natively disordered

N-terminal NORS region (1–75) alone (Fig. 1C) did also not result in significant bands for potential degradation products (Fig. 3B). His-PknG1–75 purified by Ni-affinity chromatography (Fig. 3B, lanes 5–11) was already rather pure (Fig. 3B, lanes 7–11). To further purify the protein we tried size exclusion chromatography as well as RP-HPLC chromatography. Since His-PknG1–75 is predicted to be natively disordered and since RP-HPLC under the used denaturing conditions resulted in slightly higher purity (Fig. 3B, lane 12), we will use the latter as second purification step for future purifications. Fig. 3C shows a superposition of the $^1H–^{15}N$-HSQC spectrum of $^{15}N–^{13}C$-His-PknG1–147 in black, RP-HPLC purified ^{15}N-His-PknG1–75 in red, and size exclusion chromatography (SEC) purified ^{15}N-His-PknG1–75 in green. The latter two spectra are almost identical, indicating that the used method for the second purification step for His-PknG1–75 has, as expected for a largely unfolded protein, no significant influence on the spectral appearance. Comparing the $^1H–^{15}N$-HSQC spectra of His-PknG1–147 and His-PknG1–75 (Fig. 3C) it can further be seen, that the latter represents largely a subspectrum of the first. This suggests that residues 1–75 in both PknG fragments are similarly unstructured.

Generation of a construct for the preparation of the RD region and expression and purification of the resulting PknG74–147 protein

Similar to the preparation of the expression plasmid for His-PknG1–75, we employed a site-directed mutagenesis approach to quickly obtain an expression plasmid for PknG74–147. We again used as basis the expression plasmid for His-PknG1–147, however introduced this time a factor Xa protease cleavage site (IEGR) just before residue 74 (Fig. 1C). Like His-PknG1–147 (Fig. 1C), the version including a Xa recognition site in the central region (=His-PknG1–69Xa74–147) was also highly expressed from *E. coli* BL21 (DE3) cells (Fig. 4B, lanes 1–2). After expression at 15 °C overnight, almost all of the target protein was found in the soluble fraction (Fig. 4B, lanes 1–4). As for His-PknG1–147 and His-PknG1–75, the protein purified from the centrifuged cell lysate by Ni-affinity chromatography (Fig. 4B, lanes 5–9) was already rather pure (Fig. 2B, lanes 7–9). In this case the wash fraction (Fig. 4B, lane 6) contained also some target protein, presumably to slight overloading of the used Ni-affinity column. Complete cleavage could be achieved by using a rather low factor Xa concentration and incubating overnight (~16 h, Fig. 4B, lanes 10–11). PknG74–147 was separated from the N-terminal cleaved off fragment by RP-HPLC chromatography (Fig. 4B, lane 12). Since the denaturing conditions during the RP-HPLC result in metal loss and thus unfolding, refolding was achieved by adding a reducing agent (e.g. TCEP) and slightly higher than equimolar equivalent of a divalent metal ion (e.g. Zn^{2+}) to the protein dissolved in native buffer. We prefer to use TCEP as reducing agent. Compared to dithiothreitol (DTT) it does not hamper the determination of the protein concentration by UV spectroscopy. Moreover it is rather stable under the used buffer conditions and does smell as intense as for example β-mercaptoethanol (BME). Instead of Zn ions other divalent ions maybe tried. In the crystal structure of an N-terminally truncated fragment, the cysteines of the RD coordinate a cadmium ion [6]. Fig. 4C shows a superposition of the $^1H–^{15}N$-HSQC spectrum of $^{15}N–^{13}C$-His-PknG1–147 in black and refolded ^{15}N-PknG74–147 in red. The peaks of PknG74–147 nicely overlap with the well-dispersed signals characteristic for the folded, metal-bound RD region in His-PknG1–147 purified under native conditions. This suggests that residues 74–147 in both PknG fragments adopt the same Zn-stabilized fold. By comparing the $^1H–^{15}N$-HSQC spectra of His-PknG1–147 with that of His-PknG1–75 and PknG74–147 (Figs. 3C and 4C), it becomes further evident that the spectra of the two shorter fragments add up to that observed for the

combined one. Thus the structure and dynamics of the NORS and RD region appear rather independent of each other.

Conclusion

In summary, we showed that addition of $ZnCl_2$ upon induction in minimal medium allows obtaining His-PknG1–147 with the RD in the folded, metal-bound state. We further presented a fast and efficient method to prepare expression plasmids for PknG1–147 with the native N-terminus or only the NORS (1–75) or the RD (74–147) of mycobacterial PknG (Fig. 1A) that is based on the introduction of a stop codon or a protease cleavage site in the expression plasmid for His-PknG1–147 by site directed mutagenesis (Fig. 1B and C). We further describe the expression and purification of the respective PknG proteins by Ni-affinity chromatography, followed by size exclusion chromatography or RP-HPLC chromatography as second purification step and where applicable to separate the fragments after a protease digest (Figs. 2–4B). The presented expression and purification protocols make the N-terminal regions of *M. tuberculosis* PknG amenable for a detailed structural and dynamic characterization by NMR spectroscopy (Figs. 2–4C) and other techniques and for binding studies with to be identified binding partners within the host cell as well as new potential drugs against tuberculosis.

Acknowledgments

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.pep.2015.03.015.

References

[1] J. Chao, D. Wong, X. Zheng, V. Poirier, H. Bach, Z. Hmama, Y. Av-Gay, Protein kinase and phosphatase signaling in Mycobacterium tuberculosis physiology and pathogenesis, Biochim. Biophys. Acta 2010 (1804) 620–627.

[2] Y. Av-Gay, M. Everett, The eukaryotic-like Ser/Thr protein kinases of Mycobacterium tuberculosis, Trends Microbiol. 8 (2000) 238–244.

[3] A. Walburger, A. Koul, G. Ferrari, L. Nguyen, C. Prescianotto-Baschong, K. Huygen, B. Klebl, C. Thompson, G. Bacher, J. Pieters, Protein kinase G from pathogenic mycobacteria promotes survival within macrophages, Science 304 (2004) 1800–1804.

[4] D. Tiwari, R.K. Singh, K. Goswami, S.K. Verma, B. Prakash, V.K. Nandicoori, Key residues in Mycobacterium tuberculosis protein kinase G play a role in regulating kinase activity and survival in the host, J. Biol. Chem. 284 (2009) 27467–27479.

[5] N. Scherr, P. Muller, D. Perisa, B. Combaluzier, P. Jeno, J. Pieters, Survival of pathogenic mycobacteria in macrophages is mediated through autophosphorylation of protein kinase G, J. Bacteriol. 191 (2009) 4546–4554.

[6] N. Scherr, S. Honnappa, G. Kunz, P. Mueller, R. Jayachandran, F. Winkler, J. Pieters, M.O. Steinmetz, Structural basis for the specific inhibition of protein kinase G, a virulence factor of Mycobacterium tuberculosis, Proc. Natl. Acad. Sci. U.S.A. 104 (2007) 12151–12156.

[7] M.R. Groves, D. Barford, Topological characteristics of helical repeat proteins, Curr. Opin. Struct. Biol. 9 (1999) 383–389.

[8] J. Marley, M. Lu, C. Bracken, A method for efficient isotopic labeling of recombinant proteins, J. Biomol. NMR 20 (2001) 71–75.

[9] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes, J. Biomol. NMR 6 (1995) 277–293.

[10] B.A. Johnson, Using NMRView to visualize and analyze the NMR spectra of macromolecules, Methods Mol. Biol. 278 (2004) 313–352.

[11] S.A. Dames, M. Martinez-Yamout, R.N. De Guzman, H.J. Dyson, P.E. Wright, Structural basis for Hif-1 alpha/CBP recognition in the cellular hypoxic response, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 5271–5276.

[12] O.J. Harrison, F. Bahna, P.S. Katsamba, X. Jin, J. Brasch, J. Vendome, G. Ahlsen, K.J. Carroll, S.R. Price, B. Honig, L. Shapiro, Two-step adhesive binding by classical cadherins, Nat. Struct. Mol. Biol. 17 (2010) 348–357.

Summary 2nd publication

Chemical shift assignment of the natively disordered N-terminus and the rubredoxin domain in the folded, metal bound and unfolded, oxidized state of mycobacterial protein kinase G

Matthias Wittwer, Sonja A. Dames

NMR provides valuable insights into the structural details of dynamic proteins which are inaccessible to other biophysical methods. In order to analyze residue specific dynamical properties, an NMR chemical shift assignment of the backbone signals is required. It is assumed that the N-terminal regions of mycobacterial protein kinase G (PknG) are involved in regulating the catalytic activity and sensing of the redox environment within the human host. This paper reports the chemical shift assignment of the highly flexible natively disordered N-terminal NORS region (no regulatory secondary structure) and for the redox sensitive rubredoxin domain (RD) in its folded, metal bound and its unfolded, oxidized state to the BioMagResBank. Overall, the ^{1}H, ^{13}C and ^{15}N resonance assignments for backbone residues of NORS and RD are described, as well as several sidechain ^{13}Cβ and further ^{13}C aliphatic resonances. The sidechain chemical shift assignment was limited by the intrinsic flexibility of the NORS region. Due to its unfolded character and size, the ^{1}H-^{15}N HSQC of the N-terminus contains extensive signal overlap, causing problems for further sequential resonance assignment. To resolve signal overlap in the crowded region, NMR experiments were recorded for single domain constructs (His-PknG1-75 and PknG74-147), as well as for a construct encompassing the full N-terminus (His-PknG1-147), previously described in detail [18]. Indeed, the NMR spectra of each subdomain nearly form perfect subspectra of the two-domain protein, when the RD is in the metal bound state. As expected, the NORS region shows low signal dispersion typical for an unfolded protein, whereas the redox sensitive RD in the metal bound state shows some well dispersed peaks indicative of a folded protein. Upon oxidation the RD shows crowding in the spectral random coil region implying unfolding of this motif. The presented chemical shift assignments provide the basis for the structural characterization of the regulation of PknG by redox changes and posttranslational modifications (phosphorylation).

Matthias Wittwer designed and conducted experiments, measured and carried out NMR data analysis as well as sequential assignment and was involved in writing the manuscript

CrossMark

ARTICLE

Chemical shift assignment of the intrinsically disordered N-terminus and the rubredoxin domain in the folded metal bound and unfolded oxidized state of mycobacterial protein kinase G

Matthias Wittwer[1] · Sonja A. Dames[1,2]

Abstract *Mycobacterium tuberculosis* protein kinase G (PknG) is a 82 kDa multidomain eukaryotic-like serine/threonine kinase mediating the survival of pathogenic mycobacteria within host macrophages. The N-terminal sequence preceding the catalytic kinase domain contains an approximately 75 residues long tail, which was predicted to show no regulatory secondary structure (1–75 = NORS) but harbors the major in vivo phosphorylation site (T63), and a rubredoxin-like metal binding motif (74–147 = RD). In the reduced rubredoxin motif, four conserved cysteine residues that are present as two C-X-X-C-G motifs coordinate a metal ion. The cysteines are further involved in sensing the redox environment to regulate PknG catalytic activity. Here, we report the ^1H, ^{13}C, and ^{15}N resonance assignments of the highly dynamic unstructured N-terminal region NORS and the RD in the reduced, metal bound, presumably folded and the oxidized, presumably unfolded state. Chemical shifts have been deposited at the BioMagResBank under the BMRB accession numbers 26,028 for the His-PknG1–147 with the RD in reduced, metal bound state, 26,027 for His-PknG1–75, and 26,030 and 26,029 for PknG74–147 either in the reduced, metal bound or oxidized state, respectively. The presented chemical shift assignments pave the route for the structural characterization of the regulation of PknG by redox changes and posttranslational modifications (phosphorylation).

Biological context

The mycobacterial protein kinase G (PknG) is one of 11 eukaryotic-like serine/threonine kinases encoded in *Mycobacterium tuberculosis* and related organisms (Av-Gay and Everett 2000). PknG is of particular interest regarding the development of new tuberculosis therapies, because it is secreted into infected macrophages to block the phagosome-lysosome fusion and thereby the degradation of bacteria (Walburger et al. 2004). Thus, drugs targeting PknG do not have to pass the only low permeable mycobacterial cell wall. PknG is a soluble, multidomain 82 kDa protein consisting of four different functional regions (Scherr et al. 2007). The catalytic unit is C-terminally flanked by a tetratricopeptide repeat domain (TPRD), which mediates dimerization in the crystal structure of PknG74–750 bound to the inhibitor AX20017 and may further mediate interactions with other proteins (Scherr et al. 2007). The preceding rubredoxin-like metal binding region (RD, residues ∼74–147) is characterized by two conserved C-X-X-C-G motifs. In the reduced form, the four cysteines can coordinate a divalent metal ion such as Zn^{2+} or Cd^{2+} (Scherr et al. 2007). A PknG mutant, in which the four cysteines are replaced by serines, is devoid of kinase activity and deletion of the RD renders the kinase more active against a 17mer peptide substrate derived from the in vitro substrate protein GarA (Lisa et al. 2015; Scherr et al. 2007). Moreover, PknG is more active against GarA under oxidizing conditions and mutations of the cysteines

✉ Sonja A. Dames
sonja.dames@tum.de

[1] Department of Chemistry, Biomolecular NMR Spectroscopy, Technische Universität München, Lichtenbergstr. 4, 85747 Garching, Germany

[2] Institute of Structural Biology, Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

🍃 Springer
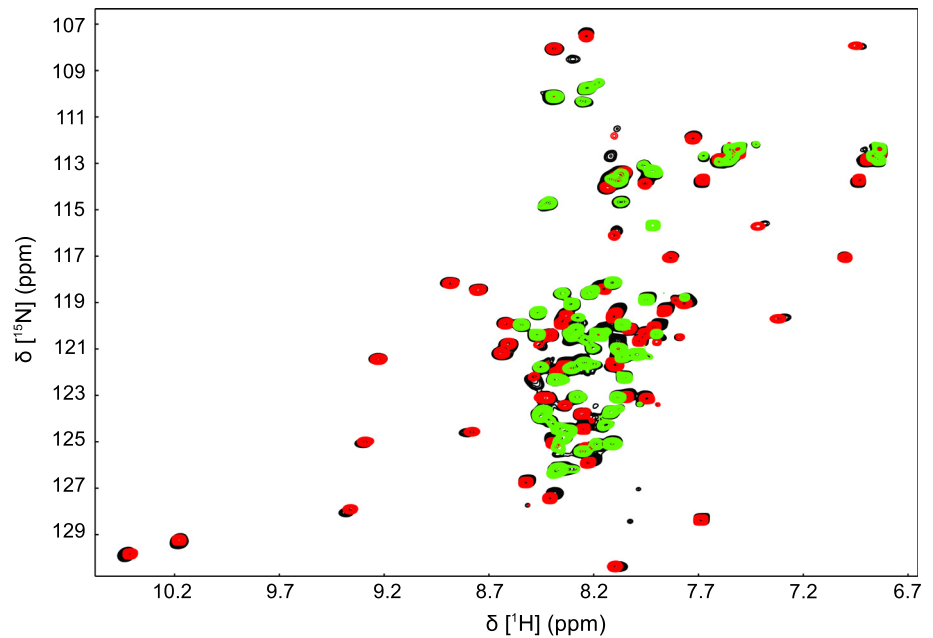
to alanines render PknG insensitive to a regulation by redox changes (Tiwari et al. 2009). The N-terminal about 75 residues were predicted to show no regulatory secondary structure (NORS) but harbor the major in vivo phosphorylation site, namely threonine 63 (Tiwari et al. 2009). Autophosphorylation in the N-terminal tail does not affect kinase activity but is required for the survival of the mycobacteria within host macrophages (Scherr et al. 2009). The structural basis of the role of the N-terminal regions for the regulation of PknG catalytic activity under different redox conditions and in different phosphorylation states is to date not well understood.

Here, we present the assignment of the resonances of the highly dynamic unstructured N-terminal region (NORS, residues ~1–75) and of the rubredoxin metal binding motif (RD, residues ~74–147) in the reduced, metal bound and the oxidized, metal free state. The assignment of the resonances provide the basis for the structural and dynamic characterization of the N-terminal regions of PknG under different redox conditions and in different phosphorylation states by NMR, which will help to better understand PknG regulation by redox changes and posttranslational modifications.

Methods and experimental

Cloning and mutagenesis

A pET-15b vector containing the coding sequence for the N-terminal regions of PknG (residues 1–147) with an N-terminal His-tag (His-PknG1–147) was kindly provided by Dr. Nicole Scherr and Prof. Dr. Jean Pieters from the Biozentrum of the University of Basel. The sequence of the 20 residue long N-terminal tag comprising 6 histidines as well as a thrombin cleavage site is: MGSSHHHHHHSSGLVPRGSH. Expression plasmids for residues 1–75 corresponding to the NORS region with an N-terminal His-tag (His-PknG1–75) and residues 74–147 corresponding to the rubredoxin domain (PknG74–147) were generated based on a mutagenesis approach by introducing either a stop codon or a factor Xa cleavage site around residue 74 (Wittwer and Dames 2015).

Expression and purification of the N-terminal regions of PknG

The expression and purification of the N-terminal constructs including His-PknG1–147, His-PknG1–75 and PknG74–147 has been described in detail earlier (Wittwer and Dames 2015). All proteins were expressed in *Escherichia coli* strain BL21 (DE3). Cells were grown in M9 minimal media supplemented with ^{15}NH$_4$Cl and/or ^{13}C-glucose as sole nitrogen or carbon sources for the preparation of ^{15}N- or ^{15}N-^{13}C-labeld proteins. His-PknG1–147 was purified in two steps based on Ni-affinity and size exclusion chromatography (Wittwer and Dames 2015). His-PknG1–75 was purified in two steps based on Ni-affinity and size exclusion chromatography or reversed-phase high performance liquid chromatography (RP-HPLC) (Wittwer and Dames 2015). PknG74–147 was purified based on Ni-affinity chromatography, followed by a factor Xa digest to cleave off His-Pkng1–73 followed by the factor Xa recognition site (IEGR) and RP-HPLC to separate the cleavage products (Wittwer and Dames 2015).

Fig. 2 ^{1}H-^{15}N-HSQC spectrum of the NORS region (His-PknG1–75) recorded at 298 K on a 500 MHz spectrometer. The sequence-specific assignments are indicated by the one-letter amino acid code and the sequence position. The additional label -sc indicates side chain amide protons. Assigned peaks of the purification tag are highlighted in blue. A asterisk indicates that the respective peaks are visible at lower contour level
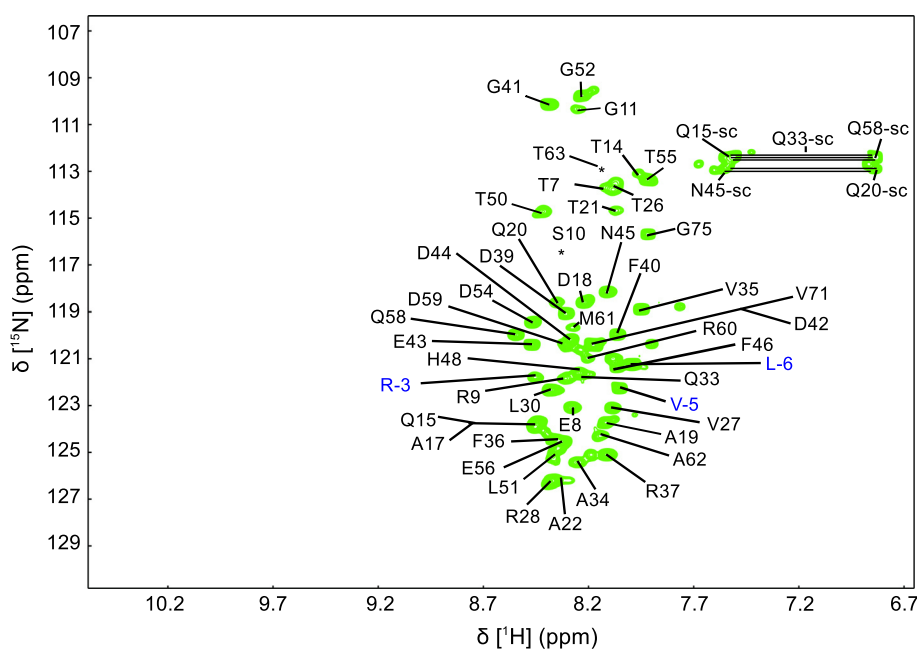
Fig. 3 ^{1}H-^{15}N-HSQC spectrum of the rubredoxin domain (PknG74–147) in the reduced, metal bound state recorded at 298 K on a 500 MHz spectrometer. The sequence-specific assignments are indicated by the one-letter amino acid code and the sequence position. The additional label -sc indicates side chain amide protons. A asterisk indicates that the respective peaks are visible at lower contour level
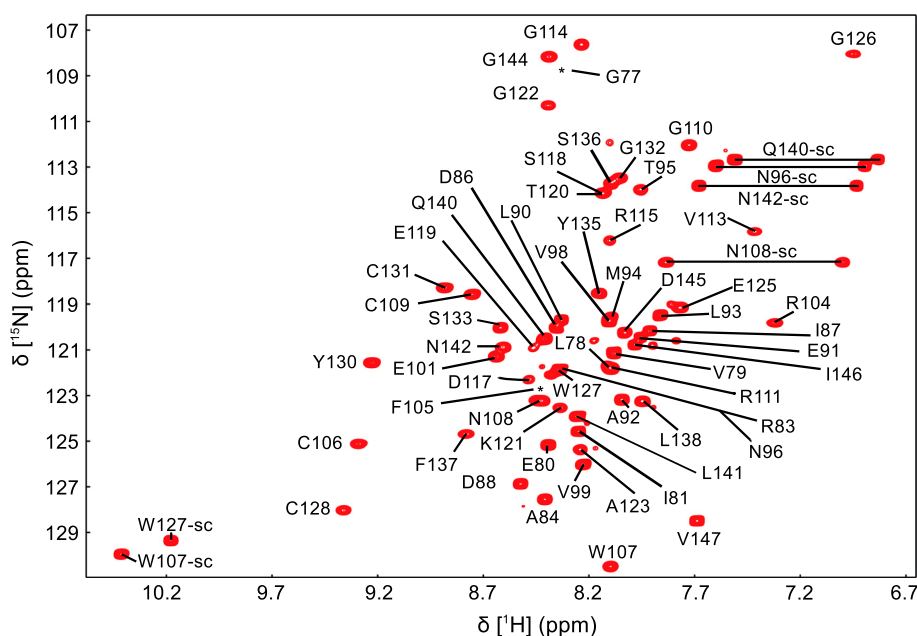
Refolding of PknG74–147 was achieved by dissolving the lyophilized protein from the final RP-HPLC step in NMR buffer (20 mM Tris, 150 mM NaCl, pH 7.5) containing the reducing agent Tris(2-carboxyethyl)phosphin (TCEP) and ZnCl$_2$ and concentrating it (Wittwer and Dames 2015).

NMR spectroscopy

NMR samples contained ^{15}N- or ^{15}N-^{13}C-labelled His-PknG1–75, His-PknG1–147 with the RD in the reduced, Zn^{2+}-bound state, or PknG74–147 in the reduced Zn^{2+}-bound or oxidized, metal free state at a concentration of

~0.4–0.8 in 20 mM Tris, 150 mM NaCl, pH 7.5, supplemented with 5 % D$_2$O and additionally 0.02–0.06 mM TCEP and 0.02–0.05 mM ZnCl$_2$ in the samples of PknG74–147 that was refolded from the lyophilized protein obtained after RP-HPLC purification (Wittwer and Dames 2015). NMR spectra were acquired at 298 K on Bruker Avance spectrometers with a magnetic field strength of 500, 600, and 900 MHz, the 500 and the 900 MHz ones equipped with cryogenic probes. The backbone assignment for ^{13}C, ^{15}N, and ^{1}H nuclei were based on three dimensional HNCA (Grzesiek and Bax 1992; Kay et al. 1994), HNCACB (Muhandiram and Kay 1994; Wittekind and

Fig. 4 ^{1}H-^{15}N-HSQC spectrum of His-PknG1–147 recorded at 298 K on a 500 MHz spectrometer. The *box* in the *upper left corner* shows a zoom of the central boxed region. The sequence-specific assignments are indicated by the one-letter amino acid code and the sequence position. The additional label *-sc* indicates side chain amide protons. Assigned peaks of the purification tag are highlighted in blue. A *asterisk* indicates that the respective peaks are visible at lower contour level
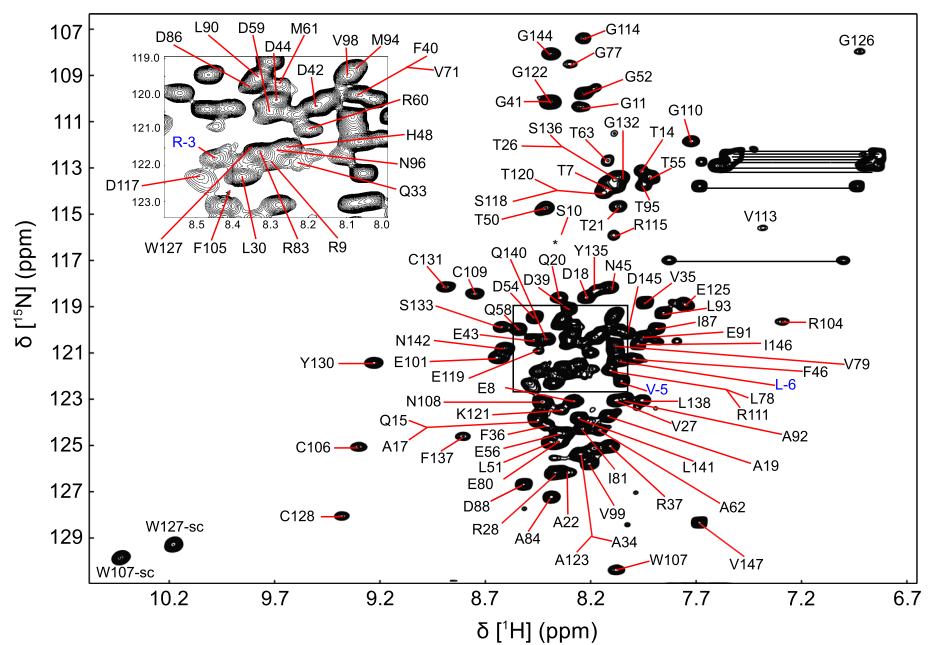
Fig. 5 Superposition of the ^{1}H-^{15}N-HSQC spectra of PknG74–147 in the oxidized, metal free (*black* recorded on a 900 MHz spectrometer) and the reduced, metal bound form (*red* recorded at 298 K on a 500 MHz spectrometer)
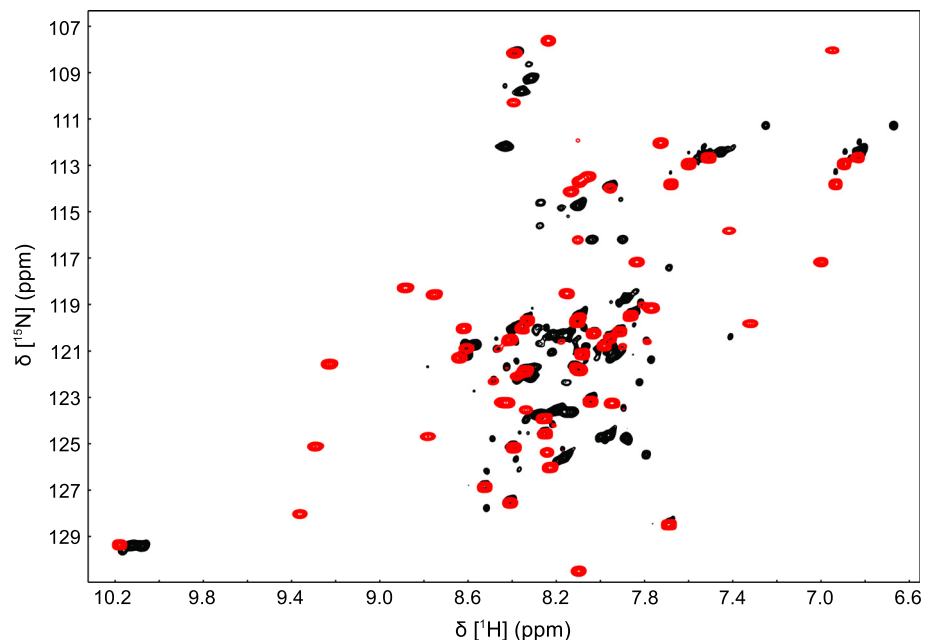
Mueller 1993), CCONH-TOCSY (Grzesiek et al. 1993; Montelione et al. 1992), HNCO (Grzesiek and Bax 1992; Kay et al. 1994), and HNHA (Vuister and Bax 1993, 1994) spectra. Data were processed with NMRPipe (Delaglio et al. 1995) and analyzed using NMRView (Johnson 2004).

Assignments and data deposition

Figure 1 shows the superposition of the ^{1}H-^{15}N-HSQC spectra of His-PknG1–147, His-PknG1–75, and PknG74–147. The spectra of each isolated region form almost perfect subspectra of that of the two region protein. This suggests that the two regions behave rather independently and do not interact significantly. This is not surprising since they are connected by a glycine-rich region (residues 75–77 = GGG). As expected the reduced, metal bound RD yields a spectrum typical for a folded protein with overall well dispersed signals, whereas the NORS region yields a spectrum typical for an intrinsically disordered protein (IDP) with a low signal resolution in the proton dimension (about 7.5–8.5 ppm). However, since the RD region also shows several signals in the crowded central region, assignment of the extensively overlapped signals of His-PknG1–147 was achieved by recording NMR assignment spectra of each functional region (NORS and RD) separately.

Fig. 6 ^{1}H-^{15}N-HSQC spectrum of the oxidized rubredoxin domain (PknG74–147) recorded at 298 K on a 900 MHz spectrometer. The sequence-specific assignments are indicated by the one-letter amino acid code and the sequence position
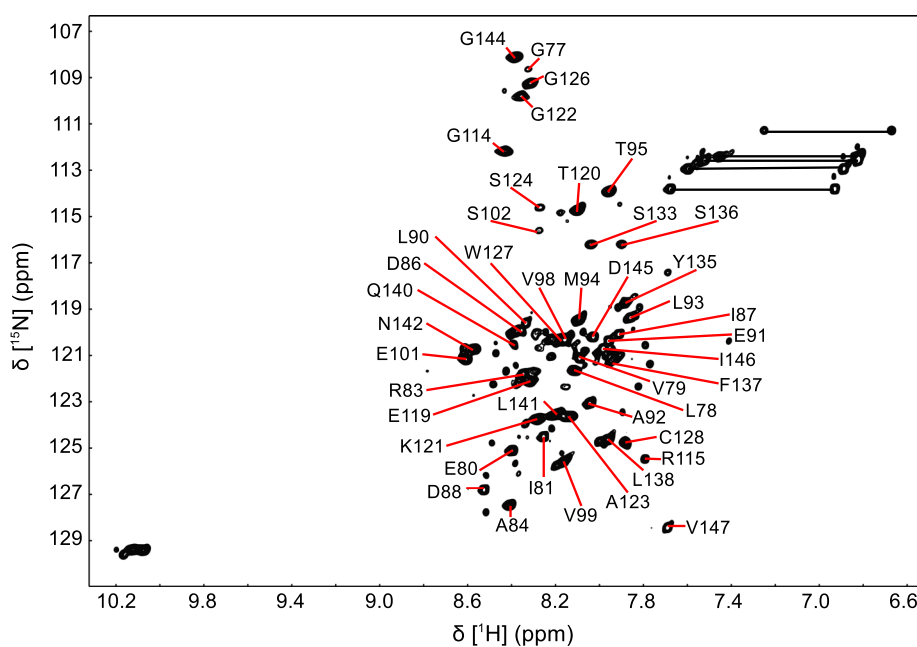
The crystal structures of PknG74–750 in complex with the inhibitor AX20017 (Scherr et al. 2007) and of PknG74–405 in complex with ADP or an ATP analogue and Mg ions (Lisa et al. 2015) have been published. Since we are mainly interested in characterizing the backbone dynamics and conformational changes induced by redox changes, phosphorylation, or interactions based on backbone RDCs as well as chemical shift changes, we focused on the assignment of the backbone atoms as well as the side chain ^{13}C$^\beta$ shifts. For the reduced, metal bound RD aliphatic side chain ^{13}C shifts have additionally been assigned. Both, the NORS and the RD region, contain each 10 prolines and several glycines (1–75: 5 G, 74–147: 9 G), which induce backbone dynamics for some residues on a time scale, resulting in broadening of the respective backbone resonances beyond detection.

For the NORS region (1–75), 72 % of the ^{1}H and ^{15}N backbone resonances of non-proline residues (65 of 75) could be assigned (Fig. 2). No chemical shift assignment could be obtained for the N-terminal residues M1-S5 and the C-terminal tail encompassing residues T64-L74, except V71 and P70. In addition, assignments were obtained for 69 % of the ^{13}C$^\alpha$, 64 % of ^{13}C$^\beta$, and 45 % of ^{13}C' resonances. The N–terminal His–tag, was only partially assigned since most resonances were not visible.

For the RD (74–147, 74 amino acids) in the reduced, metal bound form, 89 % of backbone amide ^{1}H and ^{15}N resonances of non-proline residues (64 of 74, Fig. 3), 95 % of ^{13}C$^\alpha$, 50 % of ^{1}H$^\alpha$, 97 % of ^{13}C$^\beta$, and 63 % of ^{13}C' resonances could be assigned. In this case further side chain ^{13}C resonances were obtained based on 3D CCONH–TOCSY spectra. Figure 4 shows the assigned ^{1}H-^{15}N-HSQC spectrum for the two segment protein His-PknG1–147. Chemical shift differences for the two segment compared to the single segment spectra are observed in the linker region connecting the two regions, since these residues are within the N- or C-termini in the single domain proteins but in the center of the sequence of His-Pkng1–147.

Figure 5 shows the superposition of the ^{1}H-^{15}N-HSQC spectra of the reduced, metal bound, presumably folded folded RD and the oxidized, presumably unfolded one. For the oxidized state 70 % of the backbone ^{1}H and ^{15}N resonances of non-proline residues (65 of 75, Fig. 6), 73 % of the ^{13}C$^\alpha$, and 97 % of the ^{13}C$^\beta$ could be assigned (Fig. 6). Due to the high proline (10) and glycine (5) content and the resulting dynamic behavior of the presumably unstructured oxidized rubredoxin domain, the signals for the regions surrounding the two C-X-X-C-G motifs (K103 to P112 and P129 and S133) could not be assigned.

The ^{1}H, ^{13}C, and ^{15}N chemical shifts have been deposited in the BioMagResBank under the BMRB accession numbers 26,028 for His-PknG1–147 with the RD in reduced, metal bound state, 26,027 for His-PknG1–75, and 26,030 and 26,029 for PknG74–147 either in the reduced, metal bound or oxidized state, respectively.

for his contributions while he was doing a 4-week practical and his bachelor thesis in our group. Prof. Dr. Michael Sattler and Prof. Dr. Bernd Reif from the Technische Universität München and the Helmholtz Zentrum München we thank very much for hosting our group, for continuous support and for sharing their facilities with us.

References

Av-Gay Y, Everett M (2000) The eukaryotic-like Ser/Thr protein kinases of *Mycobacterium tuberculosis*. Trends Microbiol 8:238–244

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277–293

Grzesiek S, Bax A (1992) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. J Magn Reson (1969) 96:432–440. doi:10.1016/0022-2364(92)90099-S

Grzesiek S, Anglister J, Bax A (1993) Correlation of backbone amide and aliphatic side-chain resonances in $^{13}C/^{15}N$-enriched proteins by isotropic mixing of ^{13}C magnetization. J Magn Reson Ser B 101:114–119. doi:10.1006/jmrb.1993.1019

Johnson BA (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. Methods Mol Biol 278:313–352. doi:10.1385/1-59259-809-9:313

Kay LE, Xu GY, Yamazaki T (1994) Enhanced-sensitivity triple-resonance spectroscopy with minimal H_2O saturation. J Magn Reson Ser A 109:129–133. doi:10.1006/jmra.1994.1145

Lisa MN, Gil M, Andre-Leroux G, Barilone N, Duran R, Biondi RM, Alzari PM (2015) Molecular basis of the activity and the regulation of the eukaryotic-like S/T protein kinase PknG from *Mycobacterium tuberculosis*. Structure 23:1039–1048. doi:10.1016/j.str.2015.04.001

Montelione GT, Lyons BA, Emerson SD, Tashiro M (1992) An efficient triple resonance experiment using carbon-13 isotropic mixing for determining sequence-specific resonance assignments of isotopically-enriched proteins. J Am Chem Soc 114:10974–10975. doi:10.1021/ja00053a051

Muhandiram DR, Kay LE (1994) Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. J Magn Reson Ser B 103:203–216. doi:10.1006/jmrb.1994.1032

Scherr N et al (2007) Structural basis for the specific inhibition of protein kinase G, a virulence factor of *Mycobacterium tuberculosis*. Proc Natl Acad Sci USA 104:12151–12156. doi:10.1073/pnas.0702842104

Scherr N, Muller P, Perisa D, Combaluzier B, Jeno P, Pieters J (2009) Survival of pathogenic mycobacteria in macrophages is mediated through autophosphorylation of protein kinase G. J Bacteriol 191:4546–4554. doi:10.1128/JB.00245-09

Tiwari D, Singh RK, Goswami K, Verma SK, Prakash B, Nandicoori VK (2009) Key residues in *Mycobacterium tuberculosis* protein kinase G play a role in regulating kinase activity and survival in the host. J Biol Chem 284:27467–27479. doi:10.1074/jbc.M109.036095

Vuister GW, Bax A (1993) Quantitative J correlation: a new approach for measuring homonuclear three-bond J(HNH.alpha.) coupling constants in ^{15}N-enriched proteins. J Am Chem Soc 115:7772–7777. doi:10.1021/ja00070a024

Vuister GW, Bax A (1994) Measurement of four-bond HN-Hα J-couplings in staphylococcal nuclease. J Biomol NMR 4:193–200. doi:10.1007/BF00175247

Walburger A et al (2004) Protein kinase G from pathogenic mycobacteria promotes survival within macrophages. Science 304:1800–1804. doi:10.1126/science.1099384

Wittekind M, Mueller L (1993) HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins. J Magn Reson Ser B 101:201–205. doi:10.1006/jmrb.1993.1033

Wittwer M, Dames SA (2015) Expression and purification of the natively disordered and redox sensitive metal binding regions of *Mycobacterium tuberculosis* protein kinase G. Protein Expr Purif 111:68–74. doi:10.1016/j.pep.2015.03.015

Summary 3rd publication

Oxidative Unfolding of the Rubredoxin Domain and the Natively Disordered N-terminal Region Regulate the Catalytic Activity of *Mycobacterium tuberculosis* Protein Kinase G

Matthias Wittwer, Qi Luo, Ville R. I. Kaila and Sonja A. Dames

Mycobacterium tuberculosis, the causative agent of tuberculosis, has evolved many strategies to survive human host produced reactive oxygen species (ROS) [120]. The mycobacterial protein kinase G (PknG) belongs to a group of different regulatory metal containing proteins that are sensitive to exogenous oxidative stresses and thus they are important to maintain mycobacteria in redox homeostasis [121]. In this paper, the dynamic behavior of the N-terminal regions of PknG are investigated, in order to understand how they are involved in regulating kinase activity. Dynamic properties were evaluated by NMR relaxation parameters and demonstrate distinct differences between the NORS (no regulatory secondary structure) and RD (rubredoxin domain) of the PknG N-terminus. As predicted previously, the ca. 75 residue long NORS motif is highly flexible and intrinsically unstructured, whereas the following RD in its metal-bound state is more rigid and mostly well structured. The RD adapts a folded state in a reduced, metal-bound environment and unfolds under oxidizing conditions. In addition the structure of the reduced, metal bound RD is overall similar in the absence and presence of the kinase and resemble the published crystal structures, as based on residual dipolar couplings and predicted chemical shifts. *In vitro* kinase assays lead to the assumption that the catalytic domain can phosphorylate peptide substrates better when the RD of PknG is in a folded metal bound state. However, the folding state of the substrate is of no significance when the RD of the kinase is unfolded. Thereby we assume that the substrate access to the catalytic domain is regulated by the controlled folding and unfolding of the RD. Indeed MD simulations indicate that the oxidative induced unfolding of the RD results in a more open and better accessible ATP and substrate binding cavity. Hence, redox conditions have a direct influence on PknG activity. The study illustrates the ability of PknG to respond to environmental redox changes caused by ROS within host macrophages.

Matthias Wittwer designed and conducted experiments in the wet lab, measured and carried out NMR data analysis and was involved in writing the manuscript.

Oxidative Unfolding of the Rubredoxin Domain and the Natively Disordered N-terminal Region Regulate the Catalytic Activity of *Mycobacterium tuberculosis* Protein Kinase G*[S]

Matthias Wittwer[‡], **Qi Luo**[§¶1], **Ville R. I. Kaila**[§], and ![ORCID] **Sonja A. Dames**[‡||2]

From the [‡]*Biomolecular NMR Spectroscopy and* [§]*Computational Biocatalysis, Department of Chemistry, Technische Universität München, Lichtenbergstr. 4, 85747 Garching, Germany, the* [¶]*Soft Matter Research Center and Department of Chemistry, Zhejiang University, Hangzhou 310027, China, and the* [||]*Institute of Structural Biology, Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany*

Edited by Ruma Banerjee

Mycobacterium tuberculosis escapes killing in human macrophages by secreting protein kinase G (PknG). PknG intercepts host signaling to prevent fusion of the phagosome engulfing the mycobacteria with the lysosome and, thus, their degradation. The N-terminal NORS (no regulatory secondary structure) region of PknG (approximately residues 1–75) has been shown to play a role in PknG regulation by (auto)phosphorylation, whereas the following rubredoxin-like metal-binding motif (RD, residues ~74–147) has been shown to interact tightly with the subsequent catalytic domain (approximately residues 148–420) to mediate its redox regulation. Deletions or mutations in NORS or the redox-sensitive RD significantly decrease PknG survival function. Based on combined NMR spectroscopy, *in vitro* kinase assay, and molecular dynamics simulation data, we provide novel insights into the regulatory roles of the N-terminal regions. The NORS region is indeed natively disordered and rather dynamic. Consistent with most earlier data, autophosphorylation occurs in our assays only when the NORS region is present and, thus, in the NORS region. Phosphorylation of it results only in local conformational changes and does not induce interactions with the subsequent RD. Although the reduced, metal-bound RD makes tight interactions with the following catalytic domain in the published crystal structures, it can also fold in its absence. Our data further suggest that oxidation-induced unfolding of the RD regulates substrate access to the catalytic domain and, thereby, PknG function under different redox conditions, *e.g.* when exposed to increased levels of reactive oxidative species in host macrophages.

Mycobacterium tuberculosis (Mtb),[3] the causative agent of tuberculosis, has evolved different mechanisms to monitor redox signals. The capability of Mtb to sense redox stress and to maintain redox homeostasis is important for the survival of the pathogen in the human host (1–3). Recent studies have shown that redox stress responses of mycobacteria are linked to the phosphorylation of several proteins by eukaryote-like serine/threonine kinases (4, 5). Of the 11 eukaryote-like serine/threonine kinases encoded by the Mtb genome (6), protein kinase G (PknG) and E (PknE) harbor specific redox-sensitive motifs (7, 8). The soluble PknG has been proposed to promote cellular survival of mycobacteria in host macrophages by blocking their lysosomal delivery and, thus, degradation (9). Moreover, because PknG is secreted into the cytosol of host macrophages, it is a promising drug target (9). PknG is a multidomain protein consisting of four functional regions (Fig. 1A). The N-terminal ~75 residues of the no regulatory secondary structure (NORS) region have been suggested to be intrinsically disordered and to harbor a major *in vivo* phosphorylation site at Thr[63] (8, 10, 11). Based on the crystal structure of an N-terminal truncated variant (PknG74–750) in complex with a newly detected inhibitor (AX20017, Fig. 1B), the following redox-sensitive rubredoxin-like metal binding domain (RD) makes tight interactions with the catalytic domain (8). C-terminally, the kinase domain is flanked by a tetratricopeptide repeat domain (TPRD), a structural motif typically involved in protein-protein interactions (8).

PknG can autophosphorylate itself in *trans* (10, 11). However, in contrast to other (mycobacterial) kinases, the autophosphorylation does not affect the kinase activity but is required for the survival of pathogenic mycobacteria within host macrophages (10). In rubredoxins, an iron atom is typically tetrahedrally coordinated by four cysteine residues (12), but other metal ions, such as cobalt, nickel, and zinc, can replace the iron

---

(13). The redox-sensitive RD of PknG contains two CXXCG motifs (Fig. 1A) that can, *in vitro*, coordinate a divalent metal ion, such as zinc, iron, or cadmium, in the reduced state (8, 14–16). However, it is currently unknown which metal ion is coordinated under *in vivo* conditions. Three crystal structures of PknG have been published, one of PknG74–750 (RD-KD-TPRD), with the RD coordinating Cd^{2+} and the KD in complex with the small molecule inhibitor AX20017 (PDB code 2PZI, Fig. 1B), and two of PknG74–405 (RD-KD), with the RD coordinating Zn^{2+} and the KD in complex with either an ATP analogue (ATP-γS) or ADP as well as Mg^{2+} (PDB codes 4Y12 and 4Y0X) (8, 14). The structure of the RD is very similar in all three solved structures, but the orientation of the RD with respect to the kinase domain is slightly different. In the inhibitor-bound structure, the RD interacts with the N-terminal and C-terminal lobes of the kinase domain, whereas, in the ATP analogue-bound form, the RD makes contact only with the N-terminal lobe (8, 14). It was proposed that the RD regulates the intrinsic kinase activity by restricting the accessibility of the active site (8, 14). Mutation of all four cysteines in the two conserved CXXCG motifs to alanines or serines impairs the kinase activity and renders PknG insensitive to regulation by redox changes (8, 11).

Cells of the innate immune system, such as macrophages, release high concentrations of reactive oxidative species (ROS) to kill engulfed pathogenic bacteria (17). However, based on the published crystal structures and functional data for wild-type and mutant PknG proteins, the exact mechanism of the redox regulation of the kinase activity under oxidative stress conditions remains elusive. Here we present combined NMR spectroscopy, *in vitro* kinase assay, and molecular dynamics (MD) simulation data that show how the dynamics as well as the local and global structure of PknG change upon oxidation of the RD and that the so far uncharacterized NORS region is indeed natively unfolded and the target region for PknG autophosphorylation in *trans*.

Results

The NORS Region Shows Only Local Structural Order, and the RD Can Fold in the Absence of the KD—The structural properties and dynamics of the N-terminal NORS region and of the RD in the absence of the kinase domain have not been described yet. The ^{1}H-^{15}N HSQC spectrum of the two-segment protein His-PknG1–147 (NORS-RD, supplemental Fig. S1) represents almost perfectly the sum of those of each isolated functional region (His-PknG1–75 and PknG74–147, supplemental Fig. S1). This indicates that the NORS region and the RD behave rather independently and do not interact significantly. The low signal dispersion of the NORS region indicates already that it is indeed a natively disordered protein (IDP) region. To determine in more detail the secondary structure content of each region, we measured ^{13}C$^{\alpha}$ secondary shifts (Fig. 1C) as well as ^{3}J$_{HNH\alpha}$ scalar couplings (supplemental Fig. S2A) and ^{1}H$^{\alpha}$ secondary shifts (supplemental Fig. S2B). To characterize the backbone dynamics of the two-segment protein His-PknG1–147 (NORS-RD) and of each isolated functional region (His-PknG1–75 and PknG74–147), we recorded ^{15}N relaxation data, including {^{1}H}-^{15}N NOE (Fig. 1D) as well as ^{15}N-T$_{1}$ and ^{15}N-T$_{2}$ data (supplemental Fig. S3A and Results). In addition, we back-calculated the ^{13}C$^{\alpha}$ secondary shifts from the available

crystal structures of PknG using the program Sparta+ (18) and compared them with the experimentally determined ones (supplemental Fig. S4).

The NORS region (approximately residues 1–75) shows strongly negative ^{13}C$^{\alpha}$ secondary shifts for residues preceding a proline. However, the majority of the remaining residues show only a weak propensity for α-helical secondary structure (Fig. 1C, *blue columns*). The ^{3}J$_{HNH\alpha}$ scalar couplings for the NORS region (supplemental Fig. S2A, *blue data*) are mostly between 6–8 Hz, indicating that it does not stably populate α-helical or β-sheet secondary structure and may overall only transiently and locally populate more ordered states. This is consistent with {^{1}H}-^{15}N NOE values typical for dynamic regions, with negative values below -0.2 for the isolated NORS region (Fig. 1D, *blue symbols*) and values between about 0.2 and -0.4 when connected to the RD (Fig. 1D, *black symbols*). The region between residues 20 and 40 appears overall more dynamic in His-PknG1–75 compared with His-PknG1–147, which is also reflected in the ^{15}N-T$_{1}$, and ^{15}N-T$_{2}$ data (supplemental Fig. S3A and Results). This suggests that the presence of the RD influences the dynamics of this central region of the NORS. Overall, the NORS region appears, as predicted, natively disordered. However, it may transiently populate α-helical stretches that are interrupted by several prolines (10 in total) that may induce local backbone kinks or turns (19).

The RD shows, consistent with the large chemical shift dispersion for the residues around the two CXXCG metal-coordinating motifs (^{106}CWNCG^{110} and ^{128}CPYCG^{132}), larger positive and negative ^{13}C$^{\alpha}$ secondary shifts than the NORS region (Fig. 1C, *red columns*). Thus, it appears to be present in a folded metal-bound state. Because the RD interacts tightly with the KD in the crystal structure of PknG74–750 bound to an inhibitor (PDB code 2PZI, Fig. 1B), its presence might be needed to stabilize the folded state of the RD. However, the α helix ($\sim$89–92), the 3^{10} helix ($\sim$101–103), and the two β strands ($\sim$125–127 and $\sim$134–136) detected in the crystal structure (schematically indicated at the top of Fig. 1C) appear to be similarly present in the solution state of the RD not connected to the kinase. Metal coordination to the cysteines in the two CXXCG motifs ($\sim$106–110 and $\sim$128–132) induces the formation of turns, which, as helical regions, show rather strong positive ^{13}C$^{\alpha}$ secondary shifts (Fig. 1C). The RD has as many prolines as the NORS region (each 10). Besides being involved in the already mentioned helical, sheet, and turn regions, they are further present in turns or β bridges. The region including the 3^{10} helix and the following CXXCG motifs ($\sim$100–140) shows in the isolated RD and connected to the NORS region {^{1}H}-^{15}N NOE values of $\sim$0.4–0.8 (Fig. 1D), which are typical for rather well structured regions. However, NOE values of $\sim$0–0.2 indicate that the N-terminal α-helical stretch is already more dynamic, and negative values for the N- and C-terminal ends of the isolated RD indicate even further increased backbone dynamics on the nanosecond to picosecond time scale. Because the ^{13}C$^{\alpha}$ secondary shifts and the {^{1}H}-^{15}N NOE values for the reduced, metal-bound RD are almost the same for His-PknG1–147 and PknG74–147, the presence of the NORS regions has no significant influence on its structure or backbone dynamics. This can be explained by the NORS being natively disordered as

well as connected to the RD by a dynamic linker region rich in glycines (Gly[76]-Gly[78]).

The Fold of the Isolated Reduced Metal-bound RD in Solution Is Overall Similar to the One Attached to the Kinase in Complex with AX20017, ADP, or ATP-γS in the Crystal State—Because the RD interacts tightly with the kinase domain in the available three crystal structures, albeit to a slightly different extent (8,

14), we further analyzed the effect of the presence of the kinase domain on the structure of the reduced, metal-bound state. Based on the superposition of the ^{1}H-^{15}N HSQC spectra of His-PknG74–420 encompassing the RD, the kinase, and the RD-only protein PknG74–147 (supplemental Fig. S5), peaks for several residues are found at similar chemical shift values, *e.g.* those for the two side-chain amide protons of Trp[107] and

A C[106]WNC[109]G.....18x.....C[128]PYC[131]G

1   75   147   420   750

N — NORS — RD — KD — TPRD — C   *M. tuberculosis* PknG

B

74-750 in complex with the inhibitor AX20017 and with the RD coordinating Cd[2+] (PDB-ID: 2PZI)

E

RDC [Hz]

residue sequence position

■— isolated RD-Zn[2+] (NMR)
▲— RD-Cd[2+] of 74-750 in complex with AX20017 (x-ray, PDB-ID 2PZI)
●— RD-Zn[2+] of 74-405 in complex with ADP-Mg[2+] (x-ray, PDB-ID 4Y0X)
◆— RD-Zn[2+] of 74-405 in complex with ATP-Mg[2+] (x-ray, PDB-ID 4Y0X)

C

MAKASETERSGPGTQPADAQTATSATVRPLSTQAVFRPDFGDEDNFPHPTLGPDTEPQDRMATTSRVRPPVRRLGGGLVEIPRAPDIDPLEALMNPVVPESKRFCWNCGRPVGRSDSETKGASEGWCPYCGSPYSFLPQLNPGDIV

Δδ ^{13}C^α (ppm)

α
β

D

{^{1}H}-^{15}N NOE

structured
dynamic

residue sequence position

Trp[127] of the RD. However, overall, these data did not allow us to determine in detail how similar the conformations of the RD in both proteins are. We thus recorded backbone ^{15}N-^1H RDCs to compare the conformation of the RD in the isolated state with that in the available crystal structure data. Fig. 1*E* shows a plot of the experimentally obtained RDCs as a function of the residue sequence position compared with those back-calculated from the crystal structures of His-PknG74–750 (RD-KD-TPRD) in complex with a small molecule inhibitor (PDB code 2PZI, Fig. 1*B* and supplemental Fig. S6*A*) and of PknG74–405 (RD-KD) in complex with either ADP (PDB-ID 4Y0X, supplemental Fig. S6*A*) or an ATP analogue (PDB code 4Y12, supplemental Fig. S6*A*). The respective plots of the experimental *versus* calculated values for each crystal structure and a larger representation of the RD structure are shown in supplemental Fig. S6*B*. Overall, the conformation of the isolated RD in solution is, as already indicated by comparison of the experimental and back-calculated chemical shift data (Fig. S4), very similar to that in the presence of the kinase in the three crystal structures. The observed differences around the *CXXCG* motifs (residues 106–110 and 128–132) and near residue 115 between the experimental RDCs and those back-calculated based on the crystal structures can be accounted for by missing contacts to the kinase domain in the isolated RD and slightly different contacts in PDB code 2PZI *versus* 4Y0X and 4Y12 (supplemental Fig. S6, *A* and *B*) as well as by the fact that the RD has regions with increased backbone dynamics in solution (Fig. 1*D*) and in the crystal states, *e.g.* no coordinates for residues 116–121 in PDB code 2PZI and higher B factors in the RDs of all crystal structures (supplemental Fig. S6*A*).

Deletion of the NORS and TPRD, Different Redox Conditions, and the Folding State of the Substrate Influence the Catalytic Efficiency of PknG—To better understand the relevance of the redox-sensitive RD as well as of the NORS region for the regulation of PknG function and to complement and resolve partially contradictory results from the literature (8, 10, 11, 14), we first performed *in vitro* kinase assays monitoring the progress of substrate phosphorylation based on the use of radiolabeled ATP. Because PknG can autophosphorylate itself in *trans* in the NORS region (10), we used as substrate His-PknG1–147 encompassing both the NORS and the RD (Fig. 2*A* and supple-mental Fig. S7, *A–C*). Fig. 2*A* shows the phosphorimaging kinase activity data using as kinases either His-tagged full-length WT PknG, the NORS deletion mutant His-Pkng74–750, or the NORS and TPRD deletion mutant His-PknG74–420, all with the RD in the folded, Zn-bound state, and as substrate His-PknG1–147 with the RD either in the reduced, Zn-bound or the oxidized state. Note that a fraction of the usually higher activity of the wild-type protein arises from the additional phosphorylation of its own NORS region (supplemental Fig. S7, *A–C*). Compared with the wild type (Fig. 2*A*, *blue columns*), deletion of the N-terminal NORS region reduces PknG catalytic activity significantly (Fig. 2*A*, *green columns*), and additional removal of the C-terminal TPRD results in a further reduction (Fig. 2*A*, *red columns*). This is consistent with published results (10, 11, 14). Oxidized His-PknG1–147 is overall a better substrate than the protein with the RD in the reduced, metal-bound state (Fig. 2*A*, *left versus right column* of each set of differently colored columns). This suggests that PknG autophosphorylation is more efficient when the RD of the substrate is in the oxidized state, which may facilitate binding to the catalytic cleft, especially for the phosphorylation of Thr[63] and Thr[64], which are close to the N-terminal end of the RD around residue 74. Supplemental Fig. S7*B* shows phosphorimaging kinase activity data from assays using as kinase either the full-length wild-type protein His-PknG1–750 or the NORS deletion mutant His-PknG74–750 with the RD either in the reduced, Zn-bound or the oxidized state and as substrate again His-PknG1–147 with the RD also either coordinating Zn^{2+} or being oxidized. The corresponding SDS-PAGE analysis is shown in supplemental Fig. S7*C*. The kinase with the RD in the metal-bound state phosphorylates the substrate with the RD in the oxidized state a bit better. If the RD of the kinase is in the oxidized state, then the folding state of the RD in the substrate appears to not to have a significant influence on the catalytic efficiency.

PknG Autophosphorylation Induces Only Local Conformational Changes but No Global Folding of the NORS Region—Complementary to the phosphorimaging kinase activity data, we monitored the phosphorylation of the NORS-RD protein His-PknG1–147 (Fig. 2*B*) or the NORS-only protein His-PknG1–75 (supplemental Fig. S7*D*) based on spectral changes in ^1H-^{15}N HSQC spectra. Phosphorylation of both substrates in

FIGURE 1. **The NORS region of PknG shows local structural order, and the RD can also fold in the absence of the catalytic domain.** *A*, schematic of the domain structure of PknG. PknG consists of the N-terminal NORS region (*blue*), the subsequent RD (*red*), the catalytic KD (*yellow*), and the TPRD (*green*). *B*, ribbon representation of the crystal structure of PknG74–750 (RD-KD-TPRD) in complex with the inhibitor AX20017 (*stick representation in green*, PDB code 2PZI) (8). The domain color coding is the same as in *A*. *C* and *D*, analysis of the secondary structure content and backbone dynamics of the N-terminal regions of PknG by NMR spectroscopy. *C*, $^{13}C^\alpha$ secondary shifts for His-PknG1–75 (NORS, *blue*) and PknG74–147 (RD - reduced, metal-bound, *red*) plotted as function of the amino acid sequence (49). Chemical shift values significantly higher than the random coil value indicate the presence of α-helical and those significantly lower of β-sheet secondary structure. The secondary structure elements for the RD presented above the sequence were extracted from the crystal structure of PknG74–750 in complex with AX20017 (PDB code 2PZI, see *B*) (8). Supplemental Fig. S2, *A and B* show the $^3J_{HNH\alpha}$ coupling constants and $^1H^\alpha$ secondary shifts for both isolated regions, which are both also sensitive to the secondary structure content, and supplemental Fig. S2*C* shows the $^{13}C^\alpha$ secondary shifts for His-PknG1–147. *D*, {^1H}-^{15}N NOE data for His-PknG1–75 (NORS, *red*), PknG 74–147 (RD, reduced, metal-bound, *blue*), and His-PknG1–147 (NORS-RD, reduced, metal-bound, *black*) plotted as a function of the residue sequence position. Negative to slightly positive values indicate flexible, unstructured regions, whereas positive values around 0.4–0.8 indicate well structured regions. The respective ^{15}N-T_1 as well as ^{15}N-T_2 values are shown in supplemental Fig. S3*A*. *E*, comparison of the structures of the isolated metal-bound rubredoxin-like domain in solution and in the crystal structures of PknG fragments containing, additionally, the kinase domain. The experimental residual dipolar couplings (RDCs) for the Zn^{2+}-bound RD (*black symbols*) and those back-calculated based on the three published crystal structures of PknG using the software PALES (50) were plotted as a function of the residue sequence position. The back-calculated RDCs based on the Cd^{2+}-bound RD in PknG74–750 in complex with the inhibitor AX20017 (PDB code 2PZI) (8) are represented as *green symbols*, those of the Zn^{2+}-bound RD of PknG74–405 in complex with an ATP analogue (PDB code 4Y12) (14) as *red symbols*, and those in complex with ADP (PDB code 4Y0X) (14) as *blue symbols*. A comparison of the experimental $^{13}C^\alpha$ secondary shift chemical shift values for the metal-bound RD with those back-calculated based on the published crystal structure data can be found in supplemental Fig. S4. Plots of the experimental *versus* calculated RDC values for each crystal structure and a larger representation of the RD structure as well as *ribbon representations* of the RD-KD part of each crystal structure are shown in supplemental Fig. S6. A superposition of the ^1H-^{15}N HSQC spectra of PknG74–147 and His-PknG74–420 in complex with ATP is shown in supplemental Fig. S5.
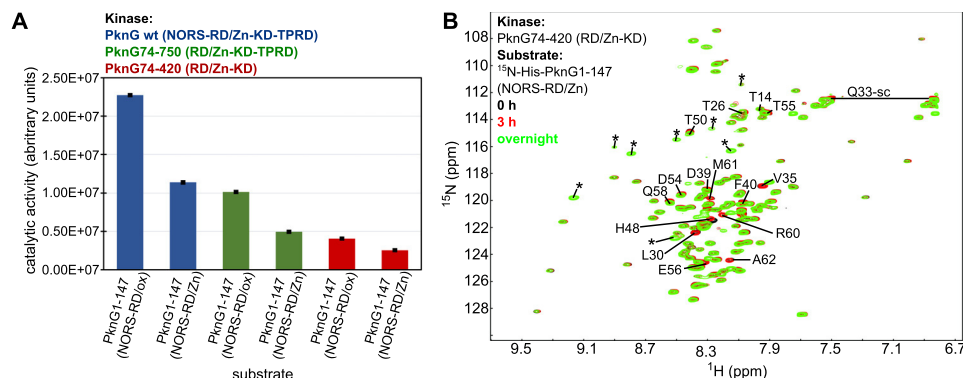
FIGURE 2. **PknG catalytic activity is sensitive to the folding state of the substrate, and autophosphorylation of the NORS region does not induce global folding.** *A*, comparison of the phosphorylation activity of His-PknG WT (*blue columns*) and the PknG deletion mutants (His-PknG74–750, green columns; His-PknG74–420, *red columns*; all with the RD in the reduced, metal-bound state (*RD/Zn*)) toward the substrate His-PknG1–147 with the RD in the reduced, metal-bound (*RD/Zn*) or oxidized state (*RD/ox*) based on phosphorimaging data. The data for the wild-type enzyme include the phosphorylation in *trans* of its own N terminus (supplemental Fig. S7A). Kinase activity data of His-PknG wild-type and of an N-terminal truncated version (His-PknG74–750) under different redox conditions and thus with the own RD in a reduced/metal-bound or oxidized state can be found in supplemental Fig. S7, *B and C. B*, NMR monitoring of *in vitro* ^{15}N-His-PknG1–147 phosphorylation by His-PknG74–420, both with the RD in the reduced metal-bound form (*RD/Zn*), based on the superposition of the ^{1}H-^{15}N HSQC spectra of unphosphorylated ^{15}N-His-PknG1–147 (*black*), and after kinase treatment for 3 h at 298 K (*red*) and further overnight at 310 K (*green*). An *asterisk* indicates a new peak that appears after phosphorylation. Assigned backbone amide cross-peaks that shift or show a change in signal intensity are labeled by the one-letter amino acid code and the sequence position (39). The additional label *-sc* indicates side chain amide protons. The NMR monitoring of the *in vitro* phosphorylation of ^{15}N-His-PknG1–75 is shown in supplemental Fig. S7D.

^{15}N-labeled form by unlabeled His-PknG74–420 (RD-KD) results in the shift or disappearance and reappearance at new positions of several peaks of the NORS region. Based on the available assignments, this includes residues near the known *in vivo* phosphorylation site Thr^{63} (11) as well as residues near other threonine residues know to be phosphorylated *in vitro* (Thr^{23}, Thr^{32}, and Thr^{64}) (10). Overall, His-PknG1–147 is more efficiently phosphorylated than His-PknG1–75. The appearance of some new more well dispersed peaks, *e.g.* around 8.7–9.3 ppm in the ^{1}H dimension (Fig. 2B and supplemental Fig. S7D), suggests that autophosphorylation induces locally more ordered states but not a global folding of the NORS region.

The PknG RD Can Switch between a Reduced, Metal-bound Folded and an Oxidized, Metal-free Unfolded State—Because replacement of the cysteines of the rubredoxin-like RD to alanines or serines disables the redox regulation of PknG kinase function, and based on fluorescence data results in significant structural changes (8, 11), we also recorded ^{1}H-^{15}N HSQC data of His-PknG1–147 (NORS-RD) under different redox conditions (Fig. 3A). When ZnCl$_2$ is added upon induction of the expression of His-PknG1–147, the black spectrum is obtained (15). This contains a subspectrum with mostly well dispersed peaks that is largely identical to the spectrum of the isolated RD (PknG74–147) that is obtained when the RP-HPLC-purified protein is refolded by adding a reducing agent, such as tris(2-carboxyethyl)phosphine, and a divalent metal ion, such as Zn^{2+} (supplemental Fig. S8A) (15). As other rubredoxin motifs, the RD can also coordinate other metal ions such as Cd^{2+}, as in the crystal structure of PknG74–750, in complex with AX20017 (8) or Mn^{2+}, Co^{2+}, or Fe^{3+}, as indicated by the ^{1}H-^{15}N HSQC data of His-PknG1–147 shown in supplemental Fig. S9. Because of the paramagnetic nature of the used metal ions, the well dispersed NMR signals of the RD are mostly not visible. However, the overall change of the spectral appearance indicates a structural transition upon metal addition (supplemental Fig. S9 and Results). Expression of PknG1–147 in minimal medium without addition of ZnCl$_2$ upon induction or addition of a rather

mild oxidizing agent such as H$_2$O$_2$ together with a metal chelator such as EDTA results in the red spectrum that shows a low dispersion for all signals. Again, the spectrum of the oxidized, metal-free isolated RD (PknG74–147, supplemental Fig. S8A; see also the {^{1}H}-^{15}N NOE data in supplemental Fig. S3B) represents a subspectrum of that of oxidized His-PknG1–147 (supplemental Fig. S8B). Indicated by the low signal dispersion, the RD in the oxidized, metal-free form is, as the NORS, largely unfolded. The spectrum of a mutant of His-PknG1–147 in which all four cysteines of the two C*XX*CG motifs have been replaced by serines (His-PknG1–147-4C/S) looks overall similar to that of the oxidized, metal-free state (supplemental Fig. S8C). Altogether, the data indicate that the RD can switch between a reduced, metal-bound folded state and an oxidized, metal-free unfolded state (Fig. 3B). Thus, a change of the redox conditions may regulate the catalytic kinase domain by controlled un- and refolding of the RD, which is expected to influence the substrate access.

MD Simulations Indicate That Oxidation of the RD Increases the Accessibility of ATP in the Substrate Binding Region—The available crystal structures for PknG fragments containing both the RD and the kinase domain do not explain how oxidation of the two C*XX*CG motifs in the RD affects the conformation and accessibility of the catalytic domain and, thus, its activity. To complement the above structural and dynamic NMR data for the RD in different redox states, we performed six independent 250-ns MD simulations of PknG74–420 (RD-KD) with the C*XX*CG motifs of the RD either coordinating a metal ion (Fe^{2+}) or with a disulfide bond in each motif (Cys^{106}-Cys^{109} and Cys^{128}-Cys^{131}) (supplemental Fig. S10A). Our simulations indicate that oxidation of the RD leads, on average, to a more open and better accessible ATP substrate-binding cavity (Fig. 3C). In line with the NMR data of the RD under different redox conditions (Fig. 3A and supplemental Fig. S8, A and B), we find that the RD containing the two C*XX*CG motifs shows overall an increased likelihood to unfold in the oxidized metal-free form, which may also favor the substrate access because of
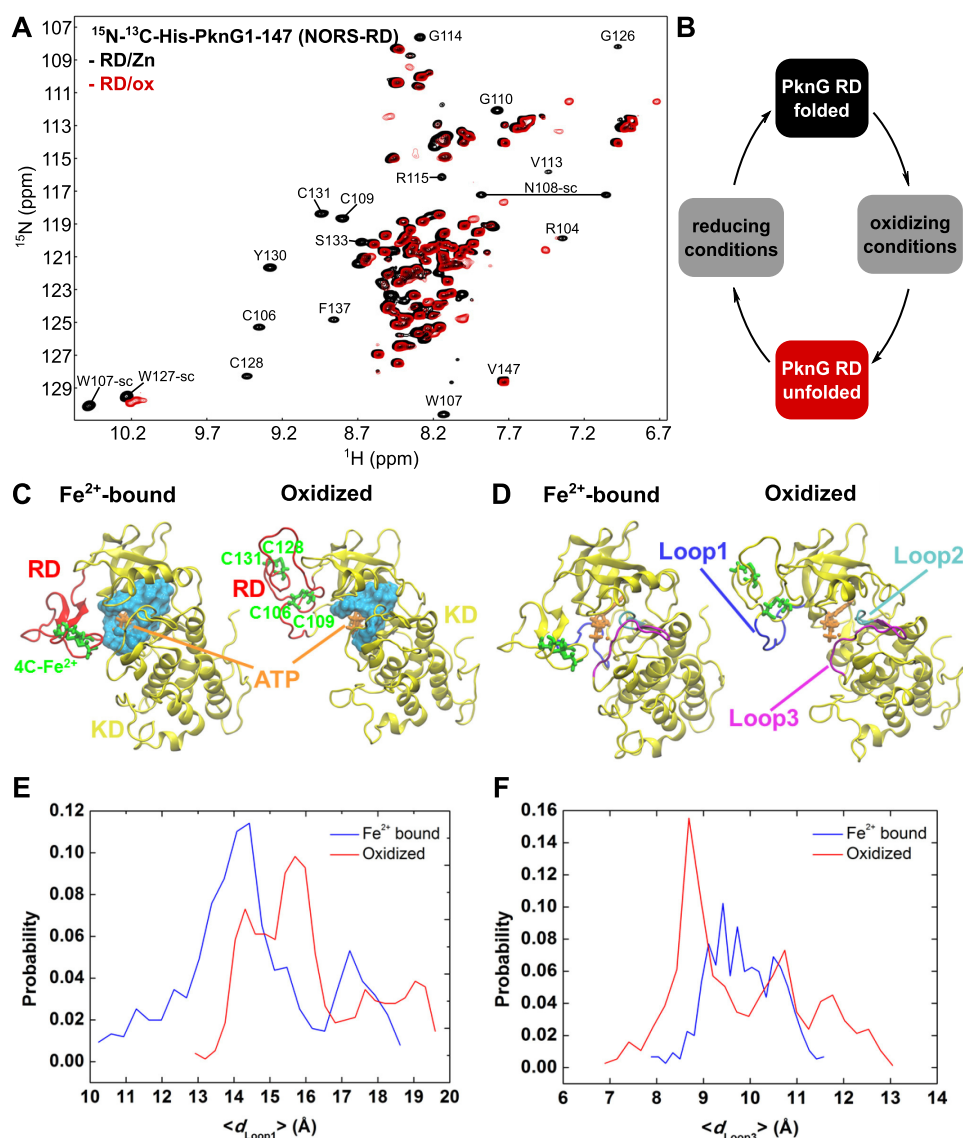
FIGURE 3. **Redox-regulated folding and unfolding of the rubredoxin-like domain of *M. tuberculosis* PknG modulates access to the catalytic cleft and, thereby, catalytic activity and/or substrate specificity.** *A*, superposition of the ^{1}H-^{15}N HSQC spectra of His-PknG1–147 in the reduced, metal-bound form (*black*) and the oxidized form (*red*). Assignments for some well resolved peaks are labeled by the one-letter amino acid code and the sequence position (39). The additional label -*sc* indicates side chain amide protons. Indicated the by the low chemical shift dispersion of all resonances under oxidizing conditions, the oxidized RD is unfolded. *B*, proposed unfolding and refolding of the RD of PknG by redox changes based on the presented NMR data in *A* and supplemental Figs. S3*B* and S8. *C–F*, characterization of conformational changes in the RD and the KD upon oxidation of the two RD C*XX*CG motifs (*green*) from classical atomistic MD simulations. *C*, the last snapshots from two 250-ns MD simulations of PknG74–420 with the cysteines of the two C*XX*CG motifs of the RD in the reduced state and coordinating Fe^{2+} (*Fe^{2+}-bound*) or in the oxidized state with disulfide bridges between Cys^{106} and Cys^{109} and Cys^{128} and Cys^{131} (*oxidized*). A similar representation for all three runs for each redox state is shown in supplemental Fig. S10*A*. The core RD region (approximately residues 100–140) is shown in *red*. Atoms within 5 Å of ATP and Mg^{2+} forming the binding cavity are shown in a space-filling representation in *cyan* to indicate that the cavity is more restricted in the Fe^{2+}-bound form. *D*, the structure of three loops surrounding the ATP-binding site, loop 1 (RD residues 94–105, *blue*), loop 2 (KD residues 292–297, *cyan*), and loop 3 (KD residues 298–310, *magenta*) in the Fe^{2+}-bound and oxidized state. *E* and *F*, the average distribution of the loop extension for loop 1 (*E*) and loop 3 (*F*) calculated from three independent 250-ns MD simulations of the Fe^{2+}-bound (*blue*) and oxidized (*red*) states. The distributions suggest that loop 1 (*E*) and loop 3 (*F*) become more extended in the oxidized state relative to the Fe^{2+}-bound state, leading to opening of the ATP-binding cavity. Loop 2 is not affected by the RD oxidization in the MD simulations (supplemental Fig. S10*B*). Plots of the volume of the ATP binding cavity (data not shown) and the SASA of residues interacting with ATP as a function of the simulation time are shown in supplemental Fig. S11, and stick representations of the region around ATP-Mg^{2+} as well as plots of the distance between Asp^{293} and Mg^{2+} as well as Asp^{276} and Lys^{278} as a function of the simulation time are shown in supplemental Fig. S12.

reduced steric clashes between the RD and the ATP binding site. To quantify the structural changes within the substrate-binding pocket, we calculated the average extension of three loops surrounding the ATP site ($<d_{Loop}>$; Fig. 3, *D–F*, and supplemental Fig. S10*B*) as well as the ATP cavity volume and solvent-accessible surface area (SASA) of residues interacting with ATP, as suggested by Scherr *et al.* (8). The simulations suggest that loops 1 and 3 surrounding the ATP-binding site

become more extended in the oxidized states (Fig. 3, *E* and *F*), making the cavity more accessible from the bulk. Oxidization of the RD also seems to lead to a somewhat larger cavity volume and SASA of ATP-surrounding residues (supplemental Fig. S11 and data not shown). In addition to these differences in substrate accessibility, our simulations further suggest that the functionally important residue Asp^{276} dissociates more from the nearby Lys^{278} in the metal-free oxidized form relative to the

metal-bound form (supplemental Fig. S12). This, in turn, might increase the kinase activity by shifting the pK_a of Asp^{276} toward higher values, thus favoring ATP hydrolysis, in line with our kinase assay data for PknG with the RD in different redox states for a more folded/bulky substrate such as PknG1–147 if the RD is in the Zn bound state (supplemental Fig. S7, *B* and *C*).

Discussion

Conformation and Dynamics of the NORS Region and PknG Autophosphorylation—Intrinsically disordered proteins (IDPs) or protein regions are typically rich in polar amino acids as well as prolines and show a high net charge (20). PknG1–75 corresponding to the NORS region contains 10 prolines and 10 negatively and seven positively charged residues as well as a high content of serines and threonines (five and 10, respectively), and, in line with this, the web program FoldIndex predicts it to be natively disordered (data not shown). Moreover, limited proteolysis of PknG resulted in a fragment lacking the first 73 residues (8). The presented NMR structural and dynamic data for the NORS region (His-PknG1–75; Fig. 1, *C* and *D*, and supplemental Figs. S2 and S3*A*) demonstrate that the NORS region is indeed rather unstructured and dynamic. Because the spectral appearance of the isolated NORS is about the same as that connected to the RD (Fig. 3*A* and supplemental Fig. S1), the two regions behave rather independently and appear not to interact. Autophosphorylation of the NORS region has been shown to play a role for the survival function of PknG (10), and some IDPs fold upon phosphorylation (20, 21). Based on the chemical shift changes of several residues near the phosphorylated threonines observed in the presented NMR monitored kinase assays using as substrates His-PknG1–147 (NORS-RD, Fig. 2*B*) or His-PknG1–75 (NORS, supplemental Fig. S7*D*), phosphorylation appears not to result in global folding but only in local structural changes, and it appears not to induce interactions with the subsequent RD. Because phosphorylation locally modulates the charge as well as conformational dynamics (20, 22), it may play a role in the interaction with regulatory proteins and/or the KD or TPRD of the same PknG molecule or a neighboring one or modulate the substrate specificity. One publication describes PknG self-cleavage that results in a fragment encompassing the NORS, the RD, and the KD and another one corresponding to the TPRD as well as autophosphorylation in the TPRD (23). However, consistent with phosphorylation data from other groups using radiolabeled ATP (10, 11, 14), our data (Fig. 2 and supplemental Fig. S7) suggest that deletion of the NORS region significantly reduces the catalytic activity and that autophosphorylation only occurs when the NORS region is present and, thus, in the NORS region.

The Role of the Redox-sensitive RD Conformation for Substrate Access to the PknG KD—In the crystal structure of PknG74–750 (RD-KD-TPRD) in complex with the small molecule inhibitor AX20017 (PDB code 2PZI, Fig. 1*B*), the RD interacts with both the N- and C-terminal lobes of the kinase domain and packs on top of the ATP binding and catalytic cleft without blocking its access (8). In the crystal structures of PknG74–405 (RD-KD) in complex with ADP-Mg^{2+} (PDB code 4Y0X, supplemental Fig. S6*A*) or ATP-γS-Mg^{2+} (PDB code 4Y12, supplemental Fig. S6*A*), the orientation of the RD is

slightly different, and it only makes contact with the N-terminal kinase lobe (14). Based on these observations, the authors suggest that the RD, in the folded metal-bound state, may regulate the catalytic activity by opening and closing the access to the substrate site (14). However, the rather small changes in the orientation of the RD relative to the kinase domain in the inhibitor-bound form and the nucleotide-bound forms appear not to be enough to explain the redox regulation of the KD by the RD. It can also not explain the inhibitory effect of AX20017 because this is a consequence of it binding to the active site. The small differences in the orientation of the RD relative to the KD in the different crystal structures arise mostly from small conformational differences because of the presence of different-size molecules in the ATP-binding region (AX20017 in 2PZI *versus* ADP-Mg^{2+} in 4Y0X or ATP-γS-Mg^{2+} in 4Y12) as well as the presence (2PZI) or absence (4Y0X and 4Y12) of the TPRD and different crystallization conditions resulting in different crystal packing. The RD has not only dynamic regions in the isolated state in solution (Fig. 1*D* and supplemental Fig. S3*A*) but also shows, in the crystal states, rather large stretches with high B factors (supplemental Fig. S6*A*). Compared with the inhibitor-bound crystal structure (2PZI), the ATP-γS-Mg^{2+}-bound (4Y12) and, even more, the ADP-Mg^{2+}-bound (4Y0X) forms appear to be overall less flexible (supplemental Fig. S6*A*). Thus, the observed variation of the orientation of the reduced, metal-bound folded RD in the three crystal structures and the observation that it contains dynamic regions in the solution and crystal states suggest that is has enough flexibility to allow binding of substrates. As described for other rubredoxin-like domains or zinc fingers (24, 25), the RD only adopts a defined three-dimensional fold upon metal binding, and the latter has not only a stabilizing effect, as, for example, in the protein IscU (26). Based on the presented NMR, MD, and kinase assay data of the RD in different redox states, the redox regulation of PknG kinase function is rather achieved by redox regulated un- and refolding of the RD, which modulates the substrate access and, thus, selectivity.

Regulation of PknG Catalytic Activity by Its Redox-sensitive RD—A redox regulation of the catalytic activity of PknG was initially suggested by kinase assay data for wild-type PknG and a mutant in which all four cysteines of the RD had been replaced by serines (PknG-C/S). Apparently, this mutant was devoid of catalytic activity toward a substrate corresponding to the N-terminal regions (His-PknG1–147), and that is phosphorylated in the natively unfolded NORS region (8, 10). Tiwari *et al.* (11) used as substrate the FHA domain containing mostly folded protein GarA, and observed the following. Deletion of the NORS and the RD together (PknG151–750) reduces the catalytic activity by ~95%, mutation of the cysteines in either C*XX*CG motif to alanines (C106A/C109A = T1, C128A/C131A = T2) by ~30%, and of both together (T1T2) by ~50–75%. Moreover, the catalytic activity of the T1T2 mutant was not very sensitive to a shift in the redox conditions (presence of 1 mM reduced or oxidized DTT), whereas wild-type PknG showed ~2.5 higher activity under oxidizing conditions (11). The latter is contradictory to the reduction of the catalytic activity by the cysteine-to-alanine replacements, which should mimic oxidizing, metal-releasing conditions. However, addi-

tion of oxidized DTT alone may not be sufficient to induce full oxidation of the two CXXCG motifs of the RD, and thus metal release and unfolding and/or the disulfide bonds result in additional local structural order (27, 28). Lisa *et al.* (14) used as substrate either GarA or only a 17-mer peptide corresponding to GarA residues 14–30. Wild-type PknG shows ~45 higher activity toward the folded GarA protein than to the GarA peptide, which, with its extended conformation, is comparable with an unfolded protein stretch. In addition, the (auto)phosphorylated NORS region has been suggested to provide phosphorylated threonine-dependent anchoring sites for high-affinity interactions with the forkhead-associated (FHA) domain of GarA (29, 30). Deletion of the N-terminal 137 residues, including the NORS region and most of the RD, resulted in higher activity toward the peptide but a bit lower toward the folded GarA protein (14). The latter is in contrast to the data by Tiwari *et al.* (11). As Lisa *et al.* already pointed out, the structural integrity of the deletion mutants has not been tested (14), and deletion of the RD may destabilize the kinase fold stronger than just oxidizing the RD CXXCG motifs. To complement the published kinase assay data and because PknG autophosphorylation has been shown to be important for mycobacterial survival (10), we used as a substrate, as Scherr *et al.* (8, 10), His-PknG1–147 and as kinase either wild type His-PknG or the deletion mutants His-PknG74–420 (RD-KD) or His-PknG74–750 (RD-KD-TPRD). Based on our data (Fig. 2 and supplemental Fig. S7), the substrate with the RD in the oxidized, unfolded state is better phosphorylated. This indicates that PknG with the RD in the reduced, metal-bound, folded state may better phosphorylate substrates with extended structures, although the kinase with the RD in the oxidized, unfolded state phosphorylates the substrate His-PknG1–147 with the RD in either redox state about equally well. Because the phosphorylated threonines are located in the mostly unstructured NORS region that adopts overall a more extended peptide-like conformation, the redox state of the RD either in the substrate or the kinase is generally expected to have an overall smaller effect than for a larger folded substrate. Interestingly, in the work of Tiwari *et al.* (11), the PknG mutant with all four cysteines of the RD CXXCG motifs replaced by alanines still showed significant activity, whereas the one with all replaced with serines in the work by Scherr *et al.* (8) showed no activity. Based on our kinase assays under different redox conditions, the results from Tiwari *et al.* (11) make more sense because mutagenesis of the cysteines should have a similar effect as oxidative unfolding. Although there are some contradictory results regarding the catalytic activity of PknG under different redox conditions and for different mutants that arise from differences in the buffer conditions (amount of substrate and kinase, "hot" and "cold" ATP, magnesium, and manganese salts, type of reducing/oxidizing agent, and so forth) and the used substrates (GarA, GarA peptide, His-PknG1–147, or kinase-dead PknG-K181M) (8, 10, 11, 14), the data altogether suggest that deletion of the RD or mutation of the cysteines in the two RD CXXCG motifs and changes in the redox conditions affect the kinase function of PknG. Based on the provided NMR and MD data about redox induced conformational changes in PknG (Fig. 3 and supplemental Figs. S3B, S8, and S10–S12) and the functional

data (Fig. 2 and Fig. S7), changes of the redox environment modulate PknG substrate selectivity more than its intrinsic catalytic efficiency.

Modulation of PknG Function by ROS in the Host Cell— The presented NMR and MD structural and dynamic data suggest that the function of PknG can be modulated by oxidative unfolding of the redox-sensitive RD, which makes the catalytic cleft more accessible for substrates (Fig. 3 and supplemental Figs. S3B and S10–S12). Similar redox-sensitive regulation mechanisms involving a four-cysteine (ZnCys$_4$) or two-cysteine two-histidine (ZnCys$_2$His$_2$) metal center have, for example, been proposed for heat shock protein 33 (Hsp33) and the mycobacterial σ factor binding protein RslA, respectively (27, 31). Other examples are the anti-σ factors RsrR (ZnCys$_3$His) and ChrR (ZnCys$_2$His$_2$) that regulate the bacterial defense against oxygen and disulfide stress (28, 32, 33). Oxidative unfolding of the PknG RD in our study has been achieved using a combination of hydrogen peroxide (H$_2$O$_2$) and a metal chelator (EDTA). The degree of oxidation and metal release depend further on other conditions such as the pH or the temperature (28). For example, for Hsp33, H$_2$O$_2$ alone is also not enough to induce oxidation and metal release. Full activation of Hsp33 requires either a combination of H$_2$O$_2$ and elevated temperatures (43 °C) or the stronger oxidant hypochlorous acid (HOCl) (27). Cells of the host innate immune system such as macrophages produce high concentrations of ROS, such as hydrogen peroxide (H$_2$O$_2$), superoxide (O$_2^-$), or HOCl and release them into the phagosome to kill engulfed pathogenic organisms (17, 34). PknG is secreted into the host cell, where it can be localized in the cytosol and the phagosome (9). Because blocking of phagosome-lysosome fusion and, thus, mycobacterial killing affords the catalytic activity of PknG, the proposed regulation of PknG activity and/or substrate specificity by oxidative unfolding of its RD makes complete sense. Future studies have to answer the question whether PknG autophosphorylation promotes mycobacterial survival in the host by just affecting its interaction with host proteins or also by phosphorylating host proteins in the cytosol and/or phagosome. One target in the host is the protein kinase C-α, a regulator of phagocytosis and the biogenesis of the phagolysosome and the closest human homologue of PknG (34, 35). PknG has been proposed to downregulate protein kinase C-α by stimulating its degradation and to be, *in vitro*, able to proteolytically cleave but not phosphorylate it (35). Future studies have to address the question whether other proteins in the host cell are targeted, which are most likely also involved in controlling phagosome-lysosome fusion, and how exactly PknG interacts with them to modulate host signaling, involving, for example, interactions of the autophosphorylated NORS region with FHA domains of human target proteins similar to those with the mycobacterial substrate protein GarA (29, 30).

Experimental Procedures

Cloning and Mutagenesis—Expression plasmids (pET-15b) for His-tagged PknG1–750 (wild-type, NORS-RD-KD-TPRD, Fig. 1A) as well as PknG1–147 (NORS-RD) and PknG74–750 (RD-KD-TPRD) were kindly provided by the group of Prof. Dr. Jean Pieters from the Biozentrum of the University of

Basel. The expression plasmids pET-15b::PknG1–75 and pET-15b::PknG74–147 were obtained by a mutagenesis-based approach as described earlier (15). The quadruple mutant His-PknG1–147-4C/4S, in which Cys[106], Cys[109], Cys[128], and Cys[131] are replaced by serines and the expression plasmid His-PknG74–420 (RD-KD) that was derived from the one for His-PknG74–750 by introducing a stop codon at position 421 were prepared using the QuikChange site-directed mutagenesis method (Stratagene, La Jolla, CA).

Protein Expression and Purification—Protein expression and purification of His-PknG1–147, His-PknG1–75, and PknG74–147 were carried out as described previously (15). All proteins were overexpressed in *Escherichia coli* BL21 (DE3) cells (Novagen). Following induction with isopropyl β-D-1-thiogalactopyranoside, cells were grown for 16 h at 15 °C, whereas, for the expression of PknG1–147-4C/4S, cells were grown for 2 h at 37 °C. For *in vitro* kinase assays, His-PknG1–750, His-PknG74–750, and His-PknG74–420 were expressed in rich medium (LB). For NMR measurements, His-PknG1–75, His-PknG1–147, and PknG74–147 were expressed in ^{15}N- or ^{15}N-^{13}C-enriched M9 minimal medium (36) supplemented with 1× BME vitamin solution (Sigma) and trace elements as described previously (15). His-PknG74–420 was expressed in ^{15}N-enriched M9 minimal medium containing 70% D$_2$O. Cells were harvested by centrifugation and sonicated, and the supernatant, after centrifugation, was loaded on a gravity flow column filled with nickel-nitrilotriacetic acid-agarose beads (Qiagen). Fractions containing significant amounts of PknG1–750, PknG74–750, or PknG74–420 were pooled, concentrated, and further purified by size exclusion chromatography using a 200 pg Superdex™ HiLoad™ 16/600 column equilibrated in 20 mM Tris and 500 mM NaCl (pH 7.5). Fractions containing the target protein were pooled and concentrated using Amicon® Ultra centrifugal filter units (MWCO 10 kDa). All purification steps were carried out at 4 °C. Following nickel affinity chromatography, fractions containing the mutant His-PknG1–147-4C/4S were loaded on a semipreparative C4 column (Jupiter®, 5 μm C4, 300 Å, 250 × 10 mm, Phenomenex) and eluted using a linear gradient from 10–90% buffer B (buffer A, H$_2$O with 0.1% TFA; buffer B, 90% acetonitrile and 10% H$_2$O with 0.1% TFA) with a flow rate of 4 ml/min within 70 min. The purified protein was lyophilized, dissolved in NMR buffer (20 mM Tris and 150 mM NaCl (pH 7.5)), washed, and concentrated using Amicon® Ultra centrifugal filter units (MWCO 10 kDa).

In Vitro Phosphorylation Observed Using a Phosphor-Imager—*In vitro* kinase assays were performed as 25-μl reactions in 20 mM Tris (pH 7.5) supplemented with 150 mM NaCl, 10 mM MgCl$_2$, 2 mM MnCl$_2$, and 40 μM [γ-^{32}P]ATP with an activity of 0.5 μCi. For the activity measurements, 0.6 μM kinase (His-PknG1–750, His-PknG74–750, His-PknG74–420) and a 5-fold molar excess of substrate (His-PknG1–147) were incubated at 30 °C for 30 min. The kinase reaction was stopped by adding 6× SDS-PAGE sample buffer and boiling the sample at 95 °C for 10 min. The kinase and/or the substrate with the RD in the oxidized form were obtained by adding a 40-fold molar excess of H$_2$O$_2$ and EDTA and incubation at 4 °C overnight. The next morning, a buffer exchange was carried out using a gravity flow Superdex™ G-25 M PD-10 column (GE Health-

care). Then the protein solution was concentrated using Amicon® Ultra centrifugal filter devices (MWCO 10 kDa) at 10,000 rpm and 4 °C. Aliquots taken during the kinase assays were separated by SDS-PAGE using 15% polyacrylamide gels. Phosphorylation of the substrate was detected by applying a phosphorimaging screen onto the gel using a Typhoon 9200 PhosphorImager. The analysis of the kinase assay data were done with the program ImageQuant (GE Healthcare).

In Vitro Phosphorylation Observed by NMR—NMR samples used to monitor substrate phosphorylation based on ^{1}H-^{15}N HSQC spectra contained 0.1 mM of ^{15}N-His-PknG1–75 or ^{15}N-His-PknG1–147 (substrate) in the presence of 1 mM ATP and 5 mM MgCl$_2$ in 20 mM Tris (pH 7.5), 150 mM NaCl, 0.05% NaN$_3$, 5% D$_2$O, and 10 μM unlabeled, catalytically active His-PknG74–420. Consecutive spectra were acquired at 298 K on a Bruker Avance 500 MHz spectrometer equipped with a cryogenic probe. Overnight incubation of the NMR sample at 310 K was done using a thermostated water bath.

NMR Sample Preparation (without Kinase Assays)—The protein concentration of the ^{15}N- and ^{15}N-^{13}C-labeled samples of His-PknG1–75, PknG74–147, and His-PknG1–147 in 20 mM Tris (pH 7.5) and 150 mM NaCl (95% H$_2$O, 5% D$_2$O) ranged from 0.1–0.8 mM. The sample of ^{15}N-^{13}C-PknG74–147 (0.2 mM) for measuring residual dipolar couplings contained ~17 mg/ml PF1 phages (ASLA Biotech). The protein concentration of ^{15}N-D-PknG74–420 in 20 mM Tris (pH 7.5), 500 mM NaCl, 10 mM tris(2-carboxyethyl)phosphine, 0.5 mM MgCl$_2$, and 0.5 mM ATP was 0.1 mM.

NMR Spectroscopy—NMR spectra were acquired at 298 K on Bruker Avance 500, 600, and 900 MHz spectrometers; the 500 and 900 MHz ones were equipped with cryogenic probes. The data were processed with NMRPipe (37) and analyzed using NMRView (38). Assignments for ^{13}C, ^{15}N, and ^{1}H nuclei were based on three-dimensional constant-time HNCA, CBCANH, CCONH-TOCSY, and HNCO spectra as described previously (39). The ^{13}C$^\alpha$ and ^{1}H$^\alpha$ secondary shifts were calculated as the difference between the measured chemical shift value and the random coil value for the respective amino acid (40). ^{3}J$_{HNH\alpha}$ coupling constants were obtained from three-dimensional HNHA spectra (41).

Information about the backbone dynamics were derived from ^{15}N relaxation data, including T$_1$ (spin-lattice relaxation), T$_2$ (spin-spin relaxation), and ^{1}H-{^{15}N} NOE. ^{15}N-^{1}H residual dipolar couplings were obtained from the analysis of ^{15}N-^{1}H-IPAP-HSQC data (42). The maximal ^{1}D$_{N-H}$ for PknG74–147 was 15.7 Hz.

Molecular Dynamics Simulations—Full atomistic molecular models of PknG74–420 with the RD in the Fe^{2+} and oxidized states were constructed based on the crystal structure of PknG74–750 in complex with the inhibitor AX20017 (PDB code 2PZI) (8). The loops missing in the crystal structure were modeled using ModLoop (43). The models were solvated in a water box with Na$^+$/Cl$^-$ ions, mimicking a 100 mM salt concentration. The molecular systems comprising ~62,500 atoms were simulated in an isothermal-isobaric (*NPT*) ensemble at $T = 310$ K and $p = 101.3$ kPa for 250 ns using 2-fs integration time steps and treating long-range electrostatic effects using the particle mesh Ewald approach. Three MD simulations for

each state, in total 1.5 μs, were performed in NAMD 2.9 (44) using the CHARMM27 force field (45) and force field parameters for the Fe^{2+}-4Cys center obtained from the literature (46). Visual Molecular Dynamics (47) was used for analyzing the MD trajectories, and cavity volumes were calculated using the *fpocket* package (48). The average extension of three loops surrounding the ATP-binding site, loop 1 (residues 94–105), loop 2 (residues 292–297), and loop 3 (residue 298–310), were calculated from the averaged Cα distances between residues $<d_{Loop1}>$: Thr[95] and Glu[101], Asn[96] and Ser[102], Pro[97] and Lys[103], and Val[98] and Arg[104]; $<d_{Loop2}>$: Ile[292] and Gly[295], Asp[293] and Ala[296], and Lys[294] and Val[297]; and $<d_{Loop3}>$: Ser[298] and Phe[303], Arg[299] and Gly[304], Ile[300] and Tyr[305], and Asn[301] and Leu[306].

References

1. Bhat, S. A., Singh, N., Trivedi, A., Kansal, P., Gupta, P., and Kumar, A. (2012) The mechanism of redox sensing in *Mycobacterium tuberculosis*. *Free Radic. Biol. Med.* **53,** 1625–1641
2. Chim, N., Johnson, P. M., and Goulding, C. W. (2014) Insights into redox sensing metalloproteins in *Mycobacterium tuberculosis*. *J. Inorg. Biochem.* **133,** 118–126
3. Kumar, A., Farhana, A., Guidry, L., Saini, V., Hondalus, M., and Steyn, A. J. (2011) Redox homeostasis in mycobacteria: the key to tuberculosis control? *Expert Rev. Mol. Med.* **13,** e39
4. Sureka, K., Hossain, T., Mukherjee, P., Chatterjee, P., Datta, P., Kundu, M., and Basu, J. (2010) Novel role of phosphorylation-dependent interaction between FtsZ and FipA in mycobacterial cell division. *PLoS ONE* **5,** e8590
5. Park, S. T., Kang, C. M., and Husson, R. N. (2008) Regulation of the SigH stress response regulon by an essential protein kinase in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 13105–13110
6. Av-Gay, Y., and Everett, M. (2000) The eukaryotic-like Ser/Thr protein kinases of *Mycobacterium tuberculosis*. *Trends Microbiol.* **8,** 238–244
7. Jayakumar, D., Jacobs, W. R., Jr, and Narayanan, S. (2008) Protein kinase E of *Mycobacterium tuberculosis* has a role in the nitric oxide stress response and apoptosis in a human macrophage model of infection. *Cell. Microbiol.* **10,** 365–374
8. Scherr, N., Honnappa, S., Kunz, G., Mueller, P., Jayachandran, R., Winkler, F., Pieters, J., and Steinmetz, M. O. (2007) Structural basis for the specific inhibition of protein kinase G, a virulence factor of *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 12151–12156
9. Walburger, A., Koul, A., Ferrari, G., Nguyen, L., Prescianotto-Baschong, C., Huygen, K., Klebl, B., Thompson, C., Bacher, G., and Pieters, J. (2004) Protein kinase G from pathogenic mycobacteria promotes survival within macrophages. *Science* **304,** 1800–1804
10. Scherr, N., Müller, P., Perisa, D., Combaluzier, B., Jenö, P., and Pieters, J. (2009) Survival of pathogenic mycobacteria in macrophages is mediated through autophosphorylation of protein kinase G. *J. Bacteriol.* **191,** 4546–4554
11. Tiwari, D., Singh, R. K., Goswami, K., Verma, S. K., Prakash, B., and Nandicoori, V. K. (2009) Key residues in *Mycobacterium tuberculosis* protein kinase G play a role in regulating kinase activity and survival in the host. *J. Biol. Chem.* **284,** 27467–27479
12. Sieker, L. C., Stenkamp, R. E., and LeGall, J. (1994) Rubredoxin in crystalline state. *Methods Enzymol.* **243,** 203–216
13. Dauter, Z., Wilson, K. S., Sieker, L. C., Moulis, J. M., and Meyer, J. (1996) Zinc- and iron-rubredoxins from *Clostridium pasteurianum* at atomic resolution: a high-precision model of a ZnS4 coordination unit in a protein. *Proc. Natl. Acad. Sci. U.S.A.* **93,** 8836–8840
14. Lisa, M. N., Gil, M., André-Leroux, G., Barilone, N., Durán, R., Biondi, R. M., and Alzari, P. M. (2015) Molecular basis of the activity and the regulation of the eukaryotic-like S/T protein kinase PknG from *Mycobacterium tuberculosis*. *Structure* **23,** 1039–1048
15. Wittwer, M., and Dames, S. A. (2015) Expression and purification of the natively disordered and redox sensitive metal binding regions of *Mycobacterium tuberculosis* protein kinase G. *Protein Expr. Purif.* **111,** 68–74
16. Gil, M., Graña, M., Schopfer, F. J., Wagner, T., Denicola, A., Freeman, B. A., Alzari, P. M., Batthyány, C., and Durán, R. (2013) Inhibition of *Mycobacterium tuberculosis* PknG by non-catalytic rubredoxin domain specific modification: reaction of an electrophilic nitro-fatty acid with the Fe-S center. *Free Radic. Biol. Med.* **65,** 150–161
17. Miller, R. A., and Britigan, B. E. (1997) Role of oxidants in microbial pathophysiology. *Clin. Microbiol. Rev.* **10,** 1–18
18. Shen, Y., and Bax, A. (2010) SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. *J. Biomol. NMR* **48,** 13–22
19. Dames, S. A., Aregger, R., Vajpai, N., Bernado, P., Blackledge, M., and Grzesiek, S. (2006) Residual dipolar couplings in short peptides reveal systematic conformational preferences of individual amino acids. *J. Am. Chem. Soc.* **128,** 13508–13514
20. Dunker, A. K., Bondos, S. E., Huang, F., and Oldfield, C. J. (2015) Intrinsically disordered proteins and multicellular organisms. *Semin. Cell Dev. Biol.* **37,** 44–55
21. Bah, A., Vernon, R. M., Siddiqui, Z., Krzeminski, M., Muhandiram, R., Zhao, C., Sonenberg, N., Kay, L. E., and Forman-Kay, J. D. (2015) Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature* **519,** 106–109
22. Tharun, I. M., Nieto, L., Haase, C., Scheepstra, M., Balk, M., Möcklinghoff, S., Adriaens, W., Dames, S. A., and Brunsveld, L. (2015) Subtype-specific modulation of estrogen receptor-coactivator interaction by phosphorylation. *ACS Chem. Biol.* **10,** 475–484
23. Cowley, S., Ko, M., Pick, N., Chow, R., Downing, K. J., Gordhan, B. G., Betts, J. C., Mizrahi, V., Smith, D. A., Stokes, R. W., and Av-Gay, Y. (2004) The *Mycobacterium tuberculosis* protein serine/threonine kinase PknG is linked to cellular glutamate/glutamine levels and is important for growth *in vivo*. *Mol. Microbiol.* **52,** 1691–1702
24. Zheng, P., Wang, Y., and Li, H. (2014) Reversible unfolding-refolding of rubredoxin: a single-molecule force spectroscopy study. *Angew. Chem. Int. Ed. Engl.* **53,** 14060–14063
25. Dames, S. A., Martinez-Yamout, M., De Guzman, R. N., Dyson, H. J., and Wright, P. E. (2002) Structural basis for Hif-1α/CBP recognition in the cellular hypoxic response. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 5271–5276
26. Iannuzzi, C., Adrover, M., Puglisi, R., Yan, R., Temussi, P. A., and Pastore, A. (2014) The role of zinc in the stability of the marginally stable IscU scaffold protein. *Protein Sci.* **23,** 1208–1219

27. Winter, J., Ilbert, M., Graf, P. C., Ozcelik, D., and Jakob, U. (2008) Bleach activates a redox-regulated chaperone by oxidative protein unfolding. *Cell* **135,** 691–701

28. Bourlès, E., Isaac, M., Lebrun, C., Latour, J. M., and Sénèque, O. (2011) Oxidation of Zn(Cys)4 zinc finger peptides by O_2 and H_2O_2: products, mechanism and kinetics. *Chemistry* **17,** 13762–13772

29. O'Hare, H. M., Durán, R., Cerveñansky, C., Bellinzoni, M., Wehenkel, A. M., Pritsch, O., Obal, G., Baumgartner, J., Vialaret, J., Johnsson, K., and Alzari, P. M. (2008) Regulation of glutamate metabolism by protein kinases in mycobacteria. *Mol. Microbiol.* **70,** 1408–1423

30. Li, J., Lee, G. I., Van Doren, S. R., and Walker, J. C. (2000) The FHA domain mediates phosphoprotein interactions. *J. Cell Sci.* **113,** 4143–4149

31. Thakur, K. G., Praveena, T., and Gopal, B. (2010) Structural and biochemical bases for the redox sensitivity of *Mycobacterium tuberculosis* RslA. *J. Mol. Biol.* **397,** 1199–1208

32. Zdanowski, K., Doughty, P., Jakimowicz, P., O'Hara, L., Buttner, M. J., Paget, M. S., and Kleanthous, C. (2006) Assignment of the zinc ligands in RsrA, a redox-sensing ZAS protein from *Streptomyces coelicolor*. *Biochemistry* **45,** 8294–8300

33. Campbell, E. A., Greenwell, R., Anthony, J. R., Wang, S., Lim, L., Das, K., Sofia, H. J., Donohue, T. J., and Darst, S. A. (2007) A conserved structural module regulates transcriptional responses to diverse stress signals in bacteria. *Mol. Cell* **27,** 793–805

34. Nguyen, L., and Pieters, J. (2005) The Trojan horse: survival tactics of pathogenic mycobacteria in macrophages. *Trends Cell Biol.* **15,** 269–276

35. Chaurasiya, S. K., and Srivastava, K. K. (2009) Downregulation of protein kinase C-α enhances intracellular survival of mycobacteria: role of PknG. *BMC Microbiol.* **9,** 271

36. Marley, J., Lu, M., and Bracken, C. (2001) A method for efficient isotopic labeling of recombinant proteins. *J. Biomol. NMR* **20,** 71–75

37. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6,** 277–293

38. Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol.* **278,** 313–352

39. Wittwer, M., and Dames, S. A. (2016) Chemical shift assignment of the intrinsically disordered N-terminus and the rubredoxin domain in the folded metal bound and unfolded oxidized state of mycobacterial protein kinase G. *Biomol. NMR Assign.* **10,** 401–406

40. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids: I: investigations of nearest-neighbor effects. *J. Biomol. NMR* **5,** 67–81

41. Vuister, G. W., and Bax, A. (1993) Quantitative J correlation: a new approach for measuring homonuclear three-bond J(HNH.α.) coupling constants in ^{15}N-enriched proteins. *J. Am. Chem. Soc.* **115,** 7772–7777

42. Ottiger, M., Delaglio, F., and Bax, A. (1998) Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.* **131,** 373–378

43. Fiser, A., and Sali, A. (2003) ModLoop: automated modeling of loops in protein structures. *Bioinformatics* **19,** 2500–2501

44. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26,** 1781–1802

45. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R., Evanseck, J., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T., Mattos, C., *et al.* (1998) a. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102,** 3586–3616

46. Gámiz-Hernández, A. P., Galstyan, A. S., and Knapp, E.-W. (2009) Understanding rubredoxin redox potentials: role of H-bonds on model complexes. *J. Chem. Theory Comput.* **5,** 2898–2908

47. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.* **14,** 33–38, 27–28

48. Le Guilloux, V., Schmidtke, P., and Tuffery, P. (2009) Fpocket: an open source platform for ligand pocket detection. *BMC Bioinformatics* **10,** 168

49. Wishart, D. S., and Sykes, B. D. (1994) The ^{13}C chemical-shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. *J. Biomol. NMR* **4,** 171–180

50. Zweckstetter, M. (2008) NMR: prediction of molecular alignment from structure using the PALES software. *Nat. Protoc.* **3,** 679–690

4. Discussion

It was shown that the N-terminal NORS region of PknG plays a role in the PknG regulation by (auto)phosphorylation, whereas the following rubredoxin-like metal bound motif interacts with the catalytic domain and is involved in redox regulation [16]. Moreover, it was shown that deletions or mutations within both domains significantly reduce the cellular survival of mycobacteria. However, based on the published crystal structures [15,16] and functional data for wild type and mutant PknG [19], the exact mechanism of the redox regulation of the kinase activity under oxidative stress conditions remains elusive.

Furthermore, no precise information about the structure and the backbone dynamics of the flexible N-terminal NORS region and the redox sensitive RD in its metal free and metal bound state is available in the current literature. In this study we investigate the structure and dynamics of both domains under different redox conditions by solution state NMR and MD simulations.

Conformation and dynamics of the N-terminal NORS region and the redox sensitive Rubredoxin domain.

In order to be able to structurally characterize the N-terminal NORS region and to elucidate the role of the redox-sensitve RD domain in redox regulation we have implemented a fast and efficient method for the preparation of several expression plasmids of the N-terminal regions. The method is based on the introduction of a stop codon or a protease cleavage site in an pET15b expression plasmid, containing the DNA sequence for the expression of PknG1-147 (NORS + RD) [15]. All constructs (His-PknG1-75, PknG74-147 & His-PknG1-147) were over expressed in BL21 E-Coli cells in a isotope enriched M9 minimal media [32] and were purified by using a protocol based on a Ni - affinity chromatography followed by SEC-(His-PknG1-147) or RP-HPLC (His-PknG1-75 & PknG74-147). The redox-sensitive RD of PknG contains two conserved CXXCG motifs, that can, *in vitro*, coordinate divalent metal ions such as zinc, iron or cadmium [14-17]. However, it is currently unknown which metal ion is coordinated under *in vivo* conditions. We could confirm by recording ^{1}H-^{15}N - HSQC that the addition of $ZnCl_2$ upon induction in M9 minimal medium allows obtaining His-PknG 1-147 with the RD in the folded, metal bound state. Furthermore, we have shown that Zn^{2+} can also be incorporated retrospectively in both RD containing constructs (His-PknG1-147 & PknG74-147) under reducing conditions.

The N-terminal NORS region was predicted to be intrinsically disordered [15]. Intrinsically disordered proteins (IDPs) or protein regions are typically rich in polar amino acids as well as prolines and show a high net charge [122]. PknG1-75 corresponding to the NORS region contains 10 prolines and 10 negatively and seven positively charged residues as well as a high content of serines and threonines, indicating that this region is rather unstructured and therefore highly flexible. Moreover, only N-terminal truncated versions of PknG could be crystallized in complex with the inhibitor AX20017 or ADP-Mg^{2+} and ATP-γS-Mg^{2+} respectively [15,16]. NMR structural and dynamic data of the NORS region demonstrate that the NORS region is indeed rather unstructured, resulting in HSQC spectra typical of IDPs with a low signal dispersion in the proton dimension (about 7.5 - 8.5 ppm). As, the spectral appearance of the isolated NORS is about the same as that connected to the RD , the two regions behave rather independently and appear not to interact. This is not surprising since they are connected by a glycine rich region (residues 75 - 77: GGG). The ^{1}H-^{15}N-HSQC data of the RD in its metal bound state show some well dispersed peaks, as expected out of the crystal structure of PknG74-750 [15].

The role of the redox sensitive RD conformation for substrate access to the kinase domain of PknG

In the crystal structure of PknG 74–750 (RD-KD-TPRD) in complex with the small molecule inhibitor AX20017 (PDB code 2PZI) the RD interacts with both the N- and C-terminal lobes of the kinase domain and packs on top of the ATP binding and catalytic cleft without blocking its access[15]. In the crystal structures of PknG 74–405 (RD-KD) in complex with ADP-Mg2_ (PDB code 4Y0X or ATP- γS Mg^{2+} (PDB code 4Y12), the orientation of the RD is slightly different, and it only makes contact with the N-terminal kinase lobe [16]. Based on these observations, the authors suggest that the RD, in the folded metal-bound state, may regulate the catalytic activity by opening and closing the access to the substrate site [16]. However, the rather small changes in the orientation of the RD relative to the kinase domain in the inhibitor - bound form and the nucleotide - bound forms appear not to be enough to explain the redox regulation of the KD by the RD. It also cannot explain the inhibitory effect of AX20017 because this is a consequence of it binding to the active site.

The small differences in the orientation of the RD relative to the KD in the different crystal structures mostly arise from small conformational differences because of the presence of different-size molecules in the ATP-binding region (AX20017 in 2PZI *versus* ADP-Mg^{2+} in 4Y0X or ATP-γS-Mg^{2+} in 4Y12) as well as the presence (2PZI) or absence (4Y0X and 4Y12) of the TPRD and different crystallization conditions resulting in different crystal packing. Not only does the RD have dynamic regions in the isolated state in solution it also shows, in the crystal states, rather large stretches with high B factors. Compared to the inhibitor bound crystal structure (2PZI), the ATP-γS-Mg^{2+} bound (4Y12) and, even more, the ADP-Mg^{2+} - bound (4Y0X) forms appear to be overall less flexible. Thus, the observed variation of the orientation of the reduced, metal bound folded RD in the three crystal structures and the observation that it contains dynamic regions in the solution and crystal states suggest that is has enough flexibility to allow the binding of substrates. As described for other rubredoxin-like domains or zinc fingers [123,124], the RD only adopts a defined three dimensional fold upon metal binding, and the latter does not only have a stabilizing effect, as, for example, in the protein IscU [125]. Based on the presented NMR, MD, and kinase assay data of the RD in different redox states, the redox regulation of PknG kinase function is rather achieved by redox regulated un- and refolding of the RD, which modulates the substrate access and, thus, selectivity.

Regulation of the PknG catalytic activity by its redox sensitive rubredoxin domain

A redox regulation of the catalytic activity of PknG was initially suggested by kinase assay data for wild-type PknG and a mutant in which all four cysteines of the RD had been replaced by serines (PknG-C/S). Apparently, this mutant was devoid of catalytic activity toward a substrate corresponding to the N-terminal regions (His-PknG1–147), and that is phosphorylated in the natively unfolded NORS region [15,23]. Tiwari *et al.* [19] used the FHA domain containing mostly folded protein GarA as a substrate, and observed the following. The deletion of the NORS and the RD together (PknG 151–750) reduces the catalytic activity by 95%, mutation of the cysteines in either CXXCG motif to alanines (C106A / C109A = T1, C128A / C131A = T2) by ~30%, and of both together (T1 T2) by ~50–75%. Moreover, the catalytic activity of the T1T2 mutant was not very sensitive to a shift in the redox conditions (presence of 1 mM reduced or oxidized DTT), whereas wild-type PknG showed ~2.5 higher activity under oxidizing conditions [19]. The latter is contradictory to the reduction of the catalytic activity by the cysteine-to-alanine replacements, which should mimic oxidizing, metal-releasing conditions.

However, the addition of oxidized DTT alone may not be sufficient to induce a full oxidation of the two CXXCG motifs of the RD, and thus metal release and unfolding and / or the disulfide bonds result in an additional local structural order [126,127]. Lisa *et al.* [16] used either GarA or only a 17-mer peptide corresponding to GarA residues 14–30 as substrate. Wild-type PknG shows ~45 higher activity toward the folded GarA protein than to the GarA peptide, which, with its extended conformation, is comparable to an unfolded protein stretch. In addition, the (auto)phosphorylated NORS region has been suggested to provide phosphorylated threonine-dependent anchoring sites for high-affinity interactions with the forkhead-associated (FHA) domain of GarA [9,128]. The deletion of the N-terminal 137 residues, including the NORS region and most of the RD, resulted in a higher activity toward the peptide but a bit lower toward the folded GarA protein [16]. The latter is in contrast to the data from Tiwari *et al.* [19]. As Lisa *et al.* already pointed out, the structural integrity of the deletion mutants has not been tested [16], and the deletion of the RD may destabilize the kinase fold stronger than just oxidizing the RD CXXCG motifs. To complement the published kinase assay data and because PknG autophosphorylation has been shown to be important for mycobacterial survival [23], we used like Scherr *et al.* [15,23], His-PknG1– 147 as a substrate and either wild type His-PknG or the deletion mutants His-PknG74–420 (RD-KD) or His-PknG74– 750 (RD-KD-TPRD) as kinase. Based on our data, the substrate with the RD in the oxidized, unfolded state is phosphorylated in a better way. This indicates that PknG with the RD in the reduced, metal-bound, folded state may phosphorylate substrates with extended structures better, although the kinase with the RD in the oxidized, unfolded state phosphorylates the substrate His-PknG1–147 with the RD in either redox state equally well. As the phosphorylated threonines are located in the mostly unstructured NORS region that overall adopts a more extended peptide-like conformation, the redox state of the RD either in the substrate or the kinase is generally expected to have an overall smaller effect than for a larger folded substrate. Interestingly, in the work of Tiwari *et al.* [19], the PknG mutant with all four cysteines of the RD CXXCG motifs replaced by alanines still showed significant activity, whereas the one with all of them replaced with serines in the work by Scherr *et al.* [15] showed no activity. Based on our kinase assays under different redox conditions, the results from Tiwari *et al.* [19] make more sense because the mutagenesis of the cysteines should have a similar effect as oxidative unfolding.

Although there are some contradictory results regarding the catalytic activity of PknG under different redox conditions and for different mutants that arise from differences in the buffer conditions (amount of substrate and kinase, "hot" and "cold" ATP, magnesium, and manganese salts, type of reducing/oxidizing agent, and so forth) and the used substrates (GarA, GarA peptide, His-PknG1–147, or kinase-dead PknG-K181M) [15,16,19,23], the data altogether suggest that the deletion of the RD or the mutation of the cysteines in the two RD CXXCG motifs and changes in the redox conditions affect the kinase function of PknG. Based on the provided NMR and MD data about redox induced conformational changes in PknG and the functional data, changes of the redox environment modulate the PknG substrate selectivity more than its intrinsic catalytic efficiency.

Modulation of PknG Function by ROS in the Host Cell

The presented NMR and MD structural and dynamic data suggest that the function of PknG can be modulated by the oxidative unfolding of the redox-sensitive RD, which makes the catalytic cleft more accessible for substrates. Similar redox-sensitive regulation mechanisms involving a four-cysteine ($ZnCys_4$) or two-cysteine two-histidine ($ZnCys_2His_2$) metal center have, for example, been proposed for heat shock protein 33 (Hsp33) and the mycobacterial _ factor binding protein RslA, respectively [126,129]. Other examples are the anti-_ factors RsrR ($ZnCys_3His$) and ChrR ($ZnCys_2His_2$) that regulate the bacterial defense against oxygen and disulfide stress [127,130,131]. The oxidative unfolding of the PknG RD in our study has been achieved using a combination of hydrogen peroxide (H_2O_2) and a metal chelator (EDTA). The degree of oxidation and metal release further depend on other conditions such as the pH or the temperature [127]. For example, for Hsp33, H_2O_2 alone also is not enough to induce oxidation and metal release. Full activation of Hsp33 requires either a combination of H_2O_2 and elevated temperatures (43 °C) or the stronger oxidant hypochlorous acid (HOCl) [126]. Cells of the host innate immune system such as macrophages produce high concentrations of ROS, such as hydrogen peroxide (H_2O_2), superoxide (O_2^-), or HOCl and release them into the phagosome to kill engulfed pathogenic organisms [132,133]. PknG is secreted into the host cell, where it can be localized in the cytosol and the phagosome [7]. Since, the blocking of phagosome-lysosome fusion and, thus, mycobacterial killing affords the catalytic activity of PknG, the proposed regulation of PknG activity and / or substrate specificity by oxidative unfolding of its RD makes complete sense.

5. Conclusions & Outlook

In this study we have focused on elucidating the structure and the dynamic behavior of the N-terminal NORS region and the rubredoxin-domain like metal binding motif (RD) by using a combination of solution state NMR spectroscopy, *in vitro* kinase assays and MD simulations in order to understand how they are involved in regulating the kinase activity of PknG. Based on experimental NMR data, we were able to confirm that the N-terminal NORS region is, as predicted previously [15], highly flexible and intrinsically unstructured, whereas the following RD in its metal-bound state is more rigid and mostly well structured. Moreover, based on *in vitro* kinase assay data we assume that the substrate access to the catalytic domain is regulated by the controlled folding and unfolding of the RD. Indeed, MD simulations indicate that the oxidative induced unfolding of the RD results in a more open and better accessible ATP and substrate binding cavity. Hence, redox conditions have a direct influence on PknG activity. Although, the study illustrates the ability of PknG to respond to environmental redox changes caused by ROS. Nevertheless, the exact mechanism of substrate binding *in vitro* and *in vivo* under different redox conditions remains unclear.

Additionally, based on residual dipolar couplings and predicted chemical shifts we have shown that the structure of the reduced, metal bound RD is overall similar in the absence and presence of the catalytic domain and resembles the published crystal structures. However, further NMR studies are required to investigate the structure and dynamics of the RD under different redox conditions in the presence of the kinase domain in liquid phase. Since this will lead to signal overcrowded NMR spectra, a segmental labeling of the RD would be a feasible solution to reduce signal overlaps and to obtain new insights into a PknG redox regulation [134].

Moreover, preliminary results (not published) have shown that the RD in its metal bound state undergoes structural changes in the presence of DPC micelles, a membrane mimetic. This may indicate that PknG can also interact with its N-terminal regions with inner membrane and / or host cell membrane proteins. Localization studies performed by Av-Gay and coworkers support this presumption [8]. Further studies, will have to investigate whether PknG mediated phosphorlylation is involved in the membrane signal transduction and what the role of the redox sensitive RD is.

Future studies will have to answer the question whether PknG autophosphorylation promotes mycobacterial survival in the host by just affecting its interaction with host proteins or also by phosphorylating host proteins in the cytosol and / or phagosome. One target in the host is the protein kinase Cα, a regulator of phagocytosis and the biogenesis of the phago-lysosome and the closest human homologue of PknG [132,135]. PknG has been proposed to downregulate protein kinase Cα by stimulating its degradation. *In vitro*, it is also able to proteolytically cleave but not phosphorylate it [135]. Furthermore, future studies will have to address the questions whether other proteins in the host cell are targeted, which of them are most likely also involved in controlling phagosome lysosome fusion, and how exactly PknG interacts with them to modulate host signaling, involving, for example, interactions of the autophosphorylated NORS region with FHA domains of human target proteins similar to those with the mycobacterial substrate protein GarA [9,128].

References

1   World Health Organization. Global Tuberculosis Report 2015. (http://apps.who.int/iris/bitstream/10665/191102/1/9789241565059_eng.pdf, Access: 26.05.2016).

2   Wallis, R. S. *et al.* Tuberculosis-advances in development of new drugs, treatment regimens, host-directed therapies, and biomarkers. *The Lancet. Infectious diseases* 16, e34-46, doi:10.1016/S1473-3099(16)00070-0 (2016).

3   Zumla, A. *et al.* Tuberculosis treatment and management--an update on treatment regimens, trials, new drugs, and adjunct therapies. *The Lancet. Respiratory medicine* 3, 220-234, doi:10.1016/S2213-2600(15)00063-6 (2015).

4   Forrellad, M. A. *et al.* Virulence factors of the Mycobacterium tuberculosis complex. *Virulence* 4, 3-66, doi:10.4161/viru.22329 (2013).

5   Warner, D. F. & Mizrahi, V. The survival kit of Mycobacterium tuberculosis. *Nature medicine* 13, 282-284, doi:10.1038/nm0307-282 (2007).

6   Parrish, N. M., Dick, J. D. & Bishai, W. R. Mechanisms of latency in Mycobacterium tuberculosis. *Trends Microbiol* 6, 107-112 (1998).

7   Walburger, A. *et al.* Protein kinase G from pathogenic mycobacteria promotes survival within macrophages. *Science* 304, 1800-1804, doi:10.1126/science.1099384 (2004).

8   Cowley, S. *et al.* The Mycobacterium tuberculosis protein serine/threonine kinase PknG is linked to cellular glutamate/glutamine levels and is important for growth in vivo. *Molecular microbiology* 52, 1691-1702, doi:10.1111/j.1365-2958.2004.04085.x (2004).

9   O'Hare, H. M. *et al.* Regulation of glutamate metabolism by protein kinases in mycobacteria. *Mol Microbiol* 70, 1408-1423, doi:10.1111/j.1365-2958.2008.06489.x (2008).

10   Ventura, M. *et al.* GarA is an essential regulator of metabolism in Mycobacterium tuberculosis. *Molecular microbiology* 90, 356-366, doi:10.1111/mmi.12368 (2013).

11   England, P. *et al.* The FHA-containing protein GarA acts as a phosphorylation-dependent molecular switch in mycobacterial signaling. *FEBS letters* 583, 301-307, doi:10.1016/j.febslet.2008.12.036 (2009).

12   Mueller, P. & Pieters, J. Identification of mycobacterial GarA as a substrate of protein kinase G from M. tuberculosis using a KESTREL-based proteome wide approach. *Journal of microbiological methods* 136, 34-39, doi:10.1016/j.mimet.2017.02.011 (2017).

13   Cole, S. T. *et al.* Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* 393, 537-544, doi:10.1038/31159 (1998).

14   Av-Gay, Y. & Everett, M. The eukaryotic-like Ser/Thr protein kinases of Mycobacterium tuberculosis. *Trends Microbiol* 8, 238-244 (2000).

15   Scherr, N. *et al.* Structural basis for the specific inhibition of protein kinase G, a virulence factor of Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences of the United States of America* 104, 12151-12156, doi:10.1073/pnas.0702842104 (2007).

16   Lisa, M. N. *et al.* Molecular Basis of the Activity and the Regulation of the Eukaryotic-like S/T Protein Kinase PknG from Mycobacterium tuberculosis. *Structure* 23, 1039-1048, doi:10.1016/j.str.2015.04.001 (2015).

17   Gil, M. *et al.* Inhibition of Mycobacterium tuberculosis PknG by non-catalytic rubredoxin domain specific modification: reaction of an electrophilic nitro-fatty acid with the Fe-S center. *Free radical biology & medicine* 65, 150-161, doi:10.1016/j.freeradbiomed.2013.06.021 (2013).

18   Wittwer, M. & Dames, S. A. Expression and purification of the natively disordered and redox sensitive metal binding regions of Mycobacterium tuberculosis protein kinase G. *Protein expression and purification* 111, 68-74, doi:10.1016/j.pep.2015.03.015 (2015)

.

19   Tiwari, D. *et al.* Key residues in Mycobacterium tuberculosis protein kinase G play a role in regulating kinase activity and survival in the host. *The Journal of biological chemistry* 284, 27467-27479, doi:10.1074/jbc.M109.036095 (2009).

20   Prisic, S. *et al.* Extensive phosphorylation with overlapping specificity by Mycobacterium tuberculosis serine/threonine protein kinases. *Proceedings of the National Academy of Sciences of the United States of America* 107, 7521-7526, doi:10.1073/pnas.0913482107 (2010).

21   Johnson, L. N., Noble, M. E. & Owen, D. J. Active and inactive protein kinases: structural basis for regulation. *Cell* 85, 149-158 (1996).

22   Prisic, S. & Husson, R. N. Mycobacterium tuberculosis Serine/Threonine Protein Kinases. *Microbiology spectrum* 2, doi:10.1128/microbiolspec.MGM2-0006-2013 (2014).

23   Scherr, N. *et al.* Survival of pathogenic mycobacteria in macrophages is mediated through autophosphorylation of protein kinase G. *Journal of bacteriology* 191, 4546-4554, doi:10.1128/JB.00245-09 (2009).

24   Wittwer, M. & Dames, S. A. Chemical shift assignment of the intrinsically disordered N-terminus and the rubredoxin domain in the folded metal bound and unfolded oxidized state of mycobacterial protein kinase G. *Biomolecular NMR assignments* 10, 401-406, doi:10.1007/s12104-016-9708-5 (2016).

25   Wittwer, M., Luo, Q., Kaila, V. R. & Dames, S. A. Oxidative Unfolding of the Rubredoxin Domain and the Natively Disordered N-terminal Region Regulate the Catalytic Activity of Mycobacterium tuberculosis Protein Kinase G. *The Journal of biological chemistry* 291, 27062-27072, doi:10.1074/jbc.M116.747089 (2016).

26   Keeler, J. *Understanding NMR Spectroscopy, 2nd Edition*. (John Wiley & Sons, 2010).

27   Levitt, M. H. *Spin Dynamics: Basics of Nuclear Magnetic Resonance*. 2nd Edition edn, (John Wiley & Sons, 2007).

28   Cavanagh, J., Fairbrother, W. J., Palmer, A. G., Skelton, N. J. & Rance, M. *Protein NMR Spectroscopy. Principles and Practice*. (Elsevier, 2006).

29   Günther, H. *NMR Spectroscopy, Third Edition* (Wiley-VCH Verlag GmbH & Co. KGaA Germany, 2013).

30   Evans, J. N. S. *Biomolecular NMR Spectroscopy*. (Oxford University Press, 1995).

31   Teng, Q. *Structural Biology: Practical NMR Applications*. (2013).

32   Marley, J., Lu, M. & Bracken, C. A method for efficient isotopic labeling of recombinant proteins. *Journal of biomolecular NMR* 20, 71-75 (2001).

33   Ardenkjaer-Larsen, J. H. *et al.* Facing and Overcoming Sensitivity Challenges in Biomolecular NMR Spectroscopy. *Angewandte Chemie* 54, 9162-9185, doi:10.1002/anie.201410653 (2015).

34   Wishart, D. S. *et al.* 1H, 13C and 15N chemical shift referencing in biomolecular NMR. *Journal of biomolecular NMR* 6, 135-140 (1995).

35   Wishart, D. S. & Sykes, B. D. Chemical shifts as a tool for structure determination. *Methods in enzymology* 239, 363-392 (1994).

36   Bloembergen, N., Purcell, E. M. & Pound, R. V. Relaxation Effects in Nuclear Magnetic Resonance Absorption. *Physical Review* 73, 679-712, doi:Doi 10.1103/Physrev.73.679 (1948).

37   Keshari, K. R. & Wilson, D. M. Chemistry and biochemistry of 13C hyperpolarized magnetic resonance using dynamic nuclear polarization. *Chemical Society reviews* 43, 1627-1659, doi:10.1039/c3cs60124b (2014).

38   Sattler, M. & Fesik, S. W. Use of deuterium labeling in NMR: overcoming a sizeable problem. *Structure (London, England : 1993)* 4, 1245-1249 (1996).

39   Gardner, K. H. & Kay, L. E. The use of 2H, 13C, 15N multidimensional NMR to study the structure and dynamics of proteins. *Annual review of biophysics and biomolecular structure* 27, 357-406, doi:10.1146/annurev.biophys.27.1.357 (1998).

40   Foster, M. P., McElroy, C. A. & Amero, C. D. Solution NMR of large molecules and assemblies. *Biochemistry* 46, 331-340, doi:10.1021/bi0621314 (2007).

41   Morris, G. A. & Freeman, R. Enhancement of nuclear magnetic resonance signals by polarization transfer. *Journal of the American Chemical Society* 101, 760-762, doi:10.1021/ja00497a058 (1979).

42   Parella, T. Pulsed field gradients: a new tool for routine NMR. *Magnetic Resonance in Chemistry* 36, 467-495, doi:10.1002/(SICI)1097-458X(199807)36:7<467::AID-OMR325>3.0.CO;2-S (1998).

43   Willker, W., Leibfritz, D., Kerssebaum, R. & Bermel, W. Gradient selection in inverse heteronuclear correlation spectroscopy. *Magnetic Resonance in Chemistry* 31, 287-292, doi:10.1002/mrc.1260310315 (2005).

44   Mandal Pravat, K. & Majumdar, A. A comprehensive discussion of HSQC and HMQC pulse sequences. *Concepts in Magnetic Resonance Part A* 20A, 1-23, doi:10.1002/cmr.a.10095 (2004).

45   Kay, L., Keifer, P. & Saarinen, T. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *Journal of the American Chemical Society* 114, 10663-10665, doi:10.1021/ja00052a088 (1992).

46   Kay, L. E., Xu, G. Y. & Yamazaki, T. Enhanced-Sensitivity Triple-Resonance Spectroscopy with Minimal H2O Saturation. *Journal of Magnetic Resonance, Series A* 109, 129-133, doi:http://dx.doi.org/10.1006/jmra.1994.1145 (1994).

47   Schleucher, J., Sattler, M. & Griesinger, C. Kohärenzselektion durch Gradienten ohne Empfindlichkeitsverlust; Anwendung auf 3D-HNCO-Experimente. *Angewandte Chemie* 105, 1518-1521, doi:10.1002/ange.19931051029 (2006).

48   Kay, L. E. Pulsed field gradient multi-dimensional NMR methods for the study of protein structure and dynamics in solution. *Progress in biophysics and molecular biology* 63, 277-299 (1995).

49   Grzesiek, S. & Bax, A. An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *Journal of Magnetic Resonance (1969)* 99, 201-207, doi:https://doi.org/10.1016/0022-2364(92)90169-8 (1992).

50   Wittekind, M. & Mueller, L. HNCACB, a High-Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha- and Beta-Carbon Resonances in Proteins. *Journal of Magnetic Resonance, Series B* 101, 201-205, doi:http://dx.doi.org/10.1006/jmrb.1993.1033 (1993).

51   Kay, L. E., Ikura, M., Tschudin, R. & Bax, A. Three-dimensional triple-resonance NMR Spectroscopy of isotopically enriched proteins. 1990. *Journal of magnetic resonance* 213, 423-441, doi:10.1016/j.jmr.2011.09.004 (1990).

52   Meissner, A. & Sorensen, O. W. Sequential HNCACB and CBCANH protein NMR pulse sequences. *Journal of magnetic resonance* 151, 328-331, doi:10.1006/jmre.2001.2374 (2001).

53   Marion, D. *et al.* Overcoming the overlap problem in the assignment of proton NMR spectra of larger proteins by use of three-dimensional heteronuclear proton-nitrogen-15 Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1.beta. *Biochemistry* 28, 6150-6156, doi:10.1021/bi00441a004 (1989).

54   Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A. & Bax, A. Three-dimensional heteronuclear NMR of nitrogen-15 labeled proteins. *Journal of the American Chemical Society* 111, 1515-1517, doi:10.1021/ja00186a066 (1989).

55   Prestegard, J. H., Bougault, C. M. & Kishore, A. I. Residual dipolar couplings in structure determination of biomolecules. *Chemical reviews* 104, 3519-3540, doi:10.1021/cr030419i (2004).

56   Tolman, J. R. & Ruan, K. NMR residual dipolar couplings as probes of biomolecular dynamics. *Chemical reviews* 106, 1720-1736, doi:10.1021/cr040429z (2006).

57   Lange, O. F. *et al.* Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science* 320, 1471-1475, doi:10.1126/science.1157092 (2008).

58   MacDonald, D. & Lu, P. Residual dipolar couplings in nucleic acid structure determination. *Current opinion in structural biology* **12, 337-343 (2002).**

59   Canales, A., Jimenez-Barbero, J. & Martin-Pastor, M. Review: Use of residual dipolar couplings to determine the structure of carbohydrates. *Magnetic resonance in chemistry : MRC* **50 Suppl 1, S80-85, doi:10.1002/mrc.3888 (2012).**

60   Tolman, J. R. Dipolar couplings as a probe of molecular dynamics and structure in solution. *Current opinion in structural biology* **11, 532-539 (2001).**

61   Bax, A., Kontaxis, G. & Tjandra, N. Dipolar couplings in macromolecular structure determination. *Methods in enzymology* **339, 127-174 (2001).**

62   Chen, K. & Tjandra, N. The use of residual dipolar coupling in studying proteins by NMR. *Topics in current chemistry* **326, 47-67, doi:10.1007/128_2011_215 (2012).**

63   Lipsitz, R. S. & Tjandra, N. Residual dipolar couplings in NMR structure analysis. *Annu Rev Bioph Biom* **33, 387-413, doi:DOI 10.1146/annurev.biophys.33.110502.140306 (2004).**

64   Brunner, E. Residual dipolar couplings in protein NMR. *Concepts in Magnetic Resonance* **13, 238-259, doi:10.1002/cmr.1012 (2001).**

65   Prestegard, J. H., al-Hashimi, H. M. & Tolman, J. R. NMR structures of biomolecules using field oriented media and residual dipolar couplings. *Quarterly reviews of biophysics* **33, 371-424 (2000).**

66   Ottiger, M., Delaglio, F. & Bax, A. Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *Journal of magnetic resonance* **131, 373-378, doi:10.1006/jmre.1998.1361 (1998).**

67   Meissner, A., Duus, J. ø. & Sørensen, O. W. Spin-State-Selective Excitation. Application for E.COSY-Type Measurement ofJHHCoupling Constants. *Journal of magnetic resonance* **128, 92-97, doi:https://doi.org/10.1006/jmre.1997.1213 (1997).**

68   Meissner, A., Duus, J. O. & Sorensen, O. W. Integration of spin-state-selective excitation into 2D NMR correlation experiments with the heteronuclear ZQ/2Q pi rotations for 1JXH-resolved E.COSY-type measurements of heteronuclear coupling constants in proteins. *Journal of biomolecular NMR* **10, 89-94 (1997).**

69   Sørensen, M. D., Meissner, A. & Sørensen, O. W. Spin-state-selective coherence transfer via intermediate states of two-spin coherence in IS spin systems: Application to E.COSY-type measurement of J coupling constants. *Journal of biomolecular NMR* **10, 181-186, doi:10.1023/A:1018323913680 (1997).**

70   Griesinger, C., Soerensen, O. W. & Ernst, R. R. Two-dimensional correlation of connected NMR transitions. *Journal of the American Chemical Society* **107, 6394-6396, doi:10.1021/ja00308a042 (1985).**

71   Pervushin, K., Riek, R., Wider, G. & Wuthrich, K. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proceedings of the National Academy of Sciences of the United States of America* **94, 12366-12371 (1997).**

72   Hu, W. & Wang, L. in *Annual Reports on NMR Spectroscopy* Vol. 58 (ed G. A. Webb) 231-303 (Academic Press, 2006).

73   Sanders, C. R. & Prestegard, J. H. Orientation and dynamics of .beta.-dodecyl glucopyranoside in phospholipid bilayers by oriented sample NMR and order matrix analysis. *Journal of the American Chemical Society* **113, 1987-1996, doi:10.1021/ja00006a019 (1991).**

74   Sanders, C. R., 2nd & Prestegard, J. H. Magnetically orientable phospholipid bilayers containing small amounts of a bile salt analogue, CHAPSO. *Biophysical journal* **58, 447-460, doi:10.1016/S0006-3495(90)82390-0 (1990).**

75   Sanders, C. R., Hare, B. J., Howard, K. P. & Prestegard, J. H. Magnetically-oriented phospholipid micelles as a tool for the study of membrane-associated molecules. *Progress in nuclear magnetic resonance spectroscopy* **26, 421-444, doi:https://doi.org/10.1016/0079-6565(94)80012-X (1994).**

76   Prestegard, J. H. & Kishore, A. I. Partial alignment of biomolecules: an aid to NMR characterization. *Current opinion in chemical biology* 5, 584-590, doi:https://doi.org/10.1016/S1367-5931(00)00247-7 (2001).

77   Tjandra, N. & Bax, A. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science* 278, 1111-1114 (1997).

78   Sanders, C. R., 2nd & Schwonek, J. P. Characterization of magnetically orientable bilayers in mixtures of dihexanoylphosphatidylcholine and dimyristoylphosphatidylcholine by solid-state NMR. *Biochemistry* 31, 8898-8905 (1992).

79   Losonczi, J. A. & Prestegard, J. H. Improved dilute bicelle solutions for high-resolution NMR of biological macromolecules. *Journal of biomolecular NMR* 12, 447-451, doi:10.1023/A:1008302110884 (1998).

80   Ramirez, B. E. & Bax, A. Modulation of the Alignment Tensor of Macromolecules Dissolved in a Dilute Liquid Crystalline Medium. *Journal of the American Chemical Society* 120, 9106-9107, doi:10.1021/ja982310b (1998).

81   Clore, G. M., Starich, M. R. & Gronenborn, A. M. Measurement of Residual Dipolar Couplings of Macromolecules Aligned in the Nematic Phase of a Colloidal Suspension of Rod-Shaped Viruses. *Journal of the American Chemical Society* 120, 10571-10572, doi:10.1021/ja982592f (1998).

82   Hansen, M. R., Mueller, L. & Pardi, A. Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nature structural biology* 5, 1065-1074, doi:10.1038/4176 (1998).

83   Zweckstetter, M. & Bax, A. Characterization of molecular alignment in aqueous suspensions of Pf1 bacteriophage. *Journal of biomolecular NMR* 20, 365-377 (2001).

84   Ottiger, M. & Bax, A. Characterization of magnetically oriented phospholipid micelles for measurement of dipolar couplings in macromolecules. *Journal of biomolecular NMR* 12, 361-372 (1998).

85   Cavagnero, S., Dyson, H. J. & Wright, P. E. Improved low pH bicelle system for orienting macromolecules over a wide temperature range. *Journal of biomolecular NMR* 13, 387-391 (1999).

86   Rückert, M. & Otting, G. Alignment of Biological Macromolecules in Novel Nonionic Liquid Crystalline Media for NMR Experiments. *Journal of the American Chemical Society* 122, 7793-7797, doi:10.1021/ja001068h (2000).

87   Koenig, B. W. *et al.* NMR Measurement of Dipolar Couplings in Proteins Aligned by Transient Binding to Purple Membrane Fragments. *Journal of the American Chemical Society* 121, 1385-1386, doi:10.1021/ja9837856 (1999).

88   Sass, H. J., Musco, G., Stahl, S. J., Wingfield, P. T. & Grzesiek, S. Solution NMR of proteins within polyacrylamide gels: diffusional properties and residual alignment by mechanical stress or embedding of oriented purple membranes. *Journal of biomolecular NMR* 18, 303-309 (2000).

89   Prosser, R. S., Losonczi, J. A. & Shiyanovskaya, I. V. Use of a Novel Aqueous Liquid Crystalline Medium for High-Resolution NMR of Macromolecules in Solution. *Journal of the American Chemical Society* 120, 11010-11011, doi:10.1021/ja982671r (1998).

90   Barrientos, L. G., Dolan, C. & Gronenborn, A. M. Characterization of surfactant liquid crystal phases suitable for molecular alignment and measurement of dipolar couplings. *Journal of biomolecular NMR* 16, 329-337 (2000).

91   Hansen, M. R., Hanson, P. & Pardi, A. Filamentous bacteriophage for aligning RNA, DNA, and proteins for measurement of nuclear magnetic resonance dipolar coupling interactions. *Methods in enzymology* 317, 220-240 (2000).

92   Tycko, R., Blanco, F. J. & Ishii, Y. Alignment of Biopolymers in Strained Gels: A New Way To Create Detectable Dipole–Dipole Couplings in High-Resolution Biomolecular NMR. *Journal of the American Chemical Society* 122, 9340-9341, doi:10.1021/ja002133q (2000).

93   Ishii, Y., Markus, M. A. & Tycko, R. Controlling residual dipolar couplings in high-resolution NMR of proteins by strain induced alignment in a gel. *Journal of biomolecular NMR* 21, 141-151 (2001).

94   Chou, J. J., Gaemers, S., Howder, B., Louis, J. M. & Bax, A. A simple apparatus for generating stretched polyacrylamide gels, yielding uniform alignment of proteins and detergent micelles. *Journal of biomolecular NMR* 21, 377-382 (2001).

95   Meier, S., Haussinger, D. & Grzesiek, S. Charged acrylamide copolymer gels as media for weak alignment. *Journal of biomolecular NMR* 24, 351-356 (2002).

96   Ruan, K. & Tolman, J. R. Composite Alignment Media for the Measurement of Independent Sets of NMR Residual Dipolar Couplings. *Journal of the American Chemical Society* 127, 15032-15033, doi:10.1021/ja055520e (2005).

97   Lorieau, J., Yao, L. & Bax, A. Liquid Crystalline Phase of G-Tetrad DNA for NMR Study of Detergent-Solubilized Proteins. *Journal of the American Chemical Society* 130, 7536-7537, doi:10.1021/ja801729f (2008).

98   Douglas, S. M., Chou, J. J. & Shih, W. M. DNA-nanotube-induced alignment of membrane proteins for NMR structure determination. *Proceedings of the National Academy of Sciences of the United States of America* 104, 6644-6648, doi:10.1073/pnas.0700930104 (2007).

99   Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR* 6, 277-293 (1995).

100   Johnson, B. A. Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods in molecular biology* 278, 313-352, doi:10.1385/1-59259-809-9:313 (2004).

101   Grzesiek, S. & Bax, A. Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *Journal of Magnetic Resonance (1969)* 96, 432-440, doi:http://dx.doi.org/10.1016/0022-2364(92)90099-S (1992).

102   Muhandiram, D. R. & Kay, L. E. Gradient-Enhanced Triple-Resonance Three-Dimensional NMR Experiments with Improved Sensitivity. *Journal of Magnetic Resonance, Series B* 103, 203-216, doi:http://dx.doi.org/10.1006/jmrb.1994.1032 (1994).

103   Grzesiek, S., Anglister, J. & Bax, A. Correlation of Backbone Amide and Aliphatic Side-Chain Resonances in 13C/15N-Enriched Proteins by Isotropic Mixing of 13C Magnetization. *Journal of Magnetic Resonance, Series B* 101, 114-119, doi:http://dx.doi.org/10.1006/jmrb.1993.1019 (1993).

104   Montelione, G. T., Lyons, B. A., Emerson, S. D. & Tashiro, M. An efficient triple resonance experiment using carbon-13 isotropic mixing for determining sequence-specific resonance assignments of isotopically-enriched proteins. *Journal of the American Chemical Society* 114, 10974-10975, doi:10.1021/ja00053a051 (1992).

105   Lyons, B. A. & Montelione, G. T. An HCCNH Triple-Resonance Experiment Using Carbon-13 Isotropic Mixing for Correlating Backbone Amide and Side-Chain Aliphatic Resonances in Isotopically Enriched Proteins. *Journal of Magnetic Resonance, Series B* 101, 206-209, doi:http://dx.doi.org/10.1006/jmrb.1993.1034 (1993).

106   Logan, T. M., Olejniczak, E. T., Xu, R. X. & Fesik, S. W. A general method for assigning NMR spectra of denatured proteins using 3D HC(CO)NH-TOCSY triple resonance experiments. *Journal of biomolecular NMR* 3, 225-231, doi:10.1007/BF00178264 (1993).

107   Clowes, R. T., Boucher, W., Hardman, C. H., Domaille, P. J. & Laue, E. D. A 4D HCC(CO)NNH experiment for the correlation of aliphatic side-chain and backbone resonances in 13C/15N-labelled proteins. *Journal of biomolecular NMR* 3, 349-354, doi:10.1007/BF00212520 (1993).

108   Carlomagno, T. *et al.* PLUSH TACSY: Homonuclear planar TACSY with two-band selective shaped pulses applied to Cα,C' transfer and Cβ,Caromatic correlations. *Journal of biomolecular NMR* 8, 161-170, doi:10.1007/BF00211162 (1996).

109   Cordier, F. & Grzesiek, S. Direct Observation of Hydrogen Bonds in Proteins by Interresidue 3hJNC' Scalar Couplings. *Journal of the American Chemical Society* 121, 1601-1602, doi:10.1021/ja983945d (1999).

110   Vuister, G. W. & Bax, A. Measurement of four-bond HN-Hα J-couplings in staphylococcal nuclease. *Journal of biomolecular NMR* 4, 193-200, doi:10.1007/BF00175247 (1994).

111   Vuister, G. W. & Bax, A. Quantitative J correlation: a new approach for measuring homonuclear three-bond J(HNH.alpha.) coupling constants in 15N-enriched proteins. *Journal of the American Chemical Society* 115, 7772-7777, doi:10.1021/ja00070a024 (1993).

112   Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S. & Sykes, B. D. 1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *Journal of biomolecular NMR* 5, 67-81 (1995).

113   Fiser, A. & Sali, A. ModLoop: automated modeling of loops in protein structures. *Bioinformatics* 19, 2500-2501 (2003).

114   Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *Journal of computational chemistry* 26, 1781-1802 (2005).

115   MacKerell, A. D. *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The journal of physical chemistry B* 102, 3586-3616 (1998).

116   Gámiz-Hernández, A. P., Galstyan, A. S. & Knapp, E.-W. Understanding rubredoxin redox potentials: role of H-bonds on model complexes. *Journal of chemical theory and computation* 5, 2898-2908 (2009).

117   Bashford, D. & Karplus, M. Multiple-site titration curves of proteins: an analysis of exact and approximate methods for their calculation. *The Journal of Physical Chemistry* 95, 9556-9561 (1991).

118   Bashford, D. & Gerwert, K. Electrostatic calculations of the pK a values of ionizable groups in bacteriorhodopsin. *Journal of molecular biology* 224, 473-486 (1992).

119   Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *Journal of molecular graphics* 14, 33-38 (1996).

120   Kumar, A. *et al.* Redox homeostasis in mycobacteria: the key to tuberculosis control? *Expert Rev Mol Med* 13, e39, doi:10.1017/S1462399411002079 (2011).

121   Chim, N., Johnson, P. M. & Goulding, C. W. Insights into redox sensing metalloproteins in Mycobacterium tuberculosis. *Journal of inorganic biochemistry* 133, 118-126, doi:10.1016/j.jinorgbio.2013.11.003 (2014).

122   Dunker, A. K., Bondos, S. E., Huang, F. & Oldfield, C. J. Intrinsically disordered proteins and multicellular organisms. *Seminars in cell & developmental biology* 37, 44-55, doi:10.1016/j.semcdb.2014.09.025 (2015).

123   Zheng, P., Wang, Y. & Li, H. Reversible unfolding-refolding of rubredoxin: a single-molecule force spectroscopy study. *Angewandte Chemie* 53, 14060-14063, doi:10.1002/anie.201408105 (2014).

124   Dames, S. A. *et al.* Residual dipolar couplings in short peptides reveal systematic conformational preferences of individual amino acids. *Journal of the American Chemical Society* 128, 13508-13514, doi:10.1021/ja063606h (2006).

125   Iannuzzi, C. *et al.* The role of zinc in the stability of the marginally stable IscU scaffold protein. *Protein science : a publication of the Protein Society* 23, 1208-1219, doi:10.1002/pro.2501 (2014).

126   Winter, J., Ilbert, M., Graf, P. C., Ozcelik, D. & Jakob, U. Bleach activates a redox-regulated chaperone by oxidative protein unfolding. *Cell* 135, 691-701, doi:10.1016/j.cell.2008.09.024 (2008).

127   Bourles, E., Isaac, M., Lebrun, C., Latour, J. M. & Seneque, O. Oxidation of Zn(Cys)4 zinc finger peptides by O2 and H2O2: products, mechanism and kinetics. *Chemistry* 17, 13762-13772, doi:10.1002/chem.201101913 (2011).

128   Li, J., Lee, G. I., Van Doren, S. R. & Walker, J. C. The FHA domain mediates phosphoprotein interactions. *Journal of Cell Science* 113, 4143 (2000).

129   Thakur, K. G., Praveena, T. & Gopal, B. Structural and biochemical bases for the redox sensitivity of Mycobacterium tuberculosis RslA. *Journal of molecular biology* 397, 1199-1208, doi:10.1016/j.jmb.2010.02.026 (2010).

130   Campbell, E. A. *et al.* A conserved structural module regulates transcriptional responses to diverse stress signals in bacteria. *Molecular cell* 27, 793-805, doi:10.1016/j.molcel.2007.07.009 (2007).

131   Zdanowski, K. *et al.* Assignment of the zinc ligands in RsrA, a redox-sensing ZAS protein from Streptomyces coelicolor. *Biochemistry* 45, 8294-8300, doi:10.1021/bi060711v (2006).

132   Nguyen, L. & Pieters, J. The Trojan horse: survival tactics of pathogenic mycobacteria in macrophages. *Trends in cell biology* 15, 269-276, doi:10.1016/j.tcb.2005.03.009 (2005).

133   Miller, R. A. & Britigan, B. E. Role of oxidants in microbial pathophysiology. *Clinical microbiology reviews* 10, 1-18 (1997).

134   Freiburger, L. *et al.* Efficient segmental isotope labeling of multi-domain proteins using Sortase A. *Journal of biomolecular NMR* 63, 1-8, doi:10.1007/s10858-015-9981-0 (2015).

135   Chaurasiya, S. K. & Srivastava, K. K. Downregulation of protein kinase C-alpha enhances intracellular survival of Mycobacteria: role of PknG. *BMC microbiology* 9, 271, doi:10.1186/1471-2180-9-271 (2009).

Appendices

<u>Supplementary part 3rd publication</u>

Oxidative unfolding of the rubredoxin domain and the natively disordered N-terminal region regulate the catalytic activity of *M. tuberculosis* protein kinase G

Matthias Wittwer[1a], Qi Luo[1b,2], Ville R. I. Kaila [1b], and Sonja A. Dames[1a,3§]

[1a/b]Technische Universität München, Department of Chemistry, [a]Biomolecular NMR Spectroscopy/[b]Computational Biocatalysis, Lichtenbergstr. 4, 85747 Garching, Germany

[2]Soft Matter Research Center and Department of Chemistry, Zhejiang University, 310027, P.R. China

[3]Institute of Structural Biology, Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

[§]To whom correspondence may be addressed. E-mail: sonja.dames@tum.de

Supplementary Information

Supplementary Results

Characterization of the backbone dynamics of the PknG N-terminal regions based on ^{15}N-relaxation data

The backbone dynamics of the isolated intrinsically disordered NORS region (His-PknG1-75), the isolated reduced, metal bound folded RD (PknG74-147), and of His-PknG1-147 containing both connected were studied by ^{15}N-relaxation experiments. The additive nature of the ^{1}H-^{15}N-HSQC data (Fig. 3A, SI Fig. S1 and S8) indicated that the NORS and the RD region behave rather independently regarding their structural properties and apparently do not interact. Based on the presented relaxation data for the two segment protein His-PknG1-147 (black symbols in Fig. 1D and SI Fig. S3A) as well as those for each region separately (blue symbols for His-PknG1-75 and red symbols PknG74-147 in Fig. 1D and SI Fig. S3A), the NORS and RD part tumble rather independently and the glycine-rich sequence (G74-G76) connecting them acts as a flexible linker. The $\{^1$H$\}$-^{15}N-NOE values of the reduced, metal bound rather well structured RD are almost the same in the isolated state (PknG74-147) and connected to the NORS region (His-PknG1-147). In the core region of the RD (residues ~100-140: ~0.4-0.8) they are typical for a folded protein. Those of the NORS region are more negative in the absence of the RD indicating greater flexibility. Whereas the unfolded NORS region represents just a rather extended chain, it represents a chain connected on one end to a roughly spherical body if connected to the folded RD. As already expected just based on the increase in molecular weight of the two segment protein His-PknG1-147 (18.001 kDa) compared to the isolated regions, His-PknG1-75 (10.241 kDa) and PknG74-147 (7.948 kDa), the presence of either region influences the ^{15}N-T$_1$ and -T$_2$ values of the other by modulating the overall tumbling. However, using a model system based on GB1 domains connected by different linker regions, it has been shown that each domain exhibits different rotational diffusion and alignment properties even if the linker was 18 residues long (1). Moreover, exchange effects arising from motions of the two PknG regions with respect to each other have to be considered. The average ^{15}N-T$_1$, -T$_2$ and $\{^1$H$\}$-^{15}N NOE values for residues 11-63 of the NORS region are 889 ± 210 ms, 267 ± 68 ms, and -0.56 ± 0.41 for the isolated form (blue symbols) and 795 ± 67 ms, 154 ± 32 ms, and -0.07 ± 0.14 if connected to the RD (black symbols), respectively. The average ^{15}N-T$_1$, -T$_2$ and $\{^1$H$\}$-^{15}N NOE values for residues 101-140 of the folded RD region are 635 ± 50 ms, 55 ± 22 ms, and 0.61 ± 0.15 for the isolated form (red symbols) and 749 ± 80 ms, 48 ± 31 ms, and 0.62 ± 0.19 if connected to the NORS region (black symbols), respectively. The ^{15}N-T$_1$ values for the isolated RD and

NORS region are about in the range expected for molecules with a size of about 8-10 kDa. The fact that they remain rather similar and do not significantly increase due the increase in molecular weight can be explained by the fact that they are flexibly linked. The ^{15}N-T_2 values for the unfolded NORS region decrease strongly due to presence of the RD, which changes the rotational diffusion properties. As mentioned above, movements of the two segments with respect to each other may result in an additional conformational exchange contribution. The ^{15}N-T_2 values for the folded RD (SI Fig. S3A) are a bit higher if connected to the NORS region. Overall they are significantly lower than expected for a completely unhindered isotropic reorientation of an ~8 kDa protein. Additional exchange effects for the RD arise from local conformational exchange processes within the RD that are also reflected in the {^{1}H}-^{15}N NOE values (Fig. 1D), e.g. residues 117-125 and that are consistent with local increased backbone dynamics indicated by higher B-factors in the available crystal structures of PknG fragments containing the RD (SI Fig. S6A). In addition the RD may experience a viscous drag due to the presence of the unfolded NORS region. A similar effect has been observed for the N-terminal domain of Formin C that contains a large unstructured loop (2).

For the isolated RD in the oxidized, unfolded state we just recorded {^{1}H}-^{15}N NOE data (SI Fig. S5B). Indicated by negative or not visible peaks in the NOE spectrum compared to the reference spectrum, this state shows NOE values that are around 0 or negative consistent with a rather unstructured and dynamic character.

Comparison of the oxidized metal free rubredoxin-like domain in the presence of different metal ions

Rubredoxins are redox sensitive metal binding proteins or protein domains, which can tetrahydrally coordinate metal ions like iron, zinc, cobalt, and cadmium by two C-X-X-C-G motifs (3). The fold and secondary structure content of the RD in the published crystal structures of protein kinase G with the rubredoxin-like domain in complex with Cd^{2+} or Zn^{2+} are very similar (4,5). SI Fig. S9 shows superpositions of the ^{1}H-^{15}N-HSQC spectra of the metal free state and different metal bound states of ^{15}N-His-PknG1-147 that confirm that the rubredoxin-like domain of PknG can also interact with various metal ions. The ^{1}H-^{15}N-HSQC spectrum of the oxidized metal free state (black) is characterized by a low signal dispersion typical for an unfolded protein. Addition of a 4-fold molar excess of a reducing agent such as TCEP and a 3 fold molar excess of Mn^{2+}, Co^{2+}, or Fe^{3+} ions results in the disappearance or shift of several resonances. Compared to the spectra of the Zn^{2+} form, fewer well dispersed peaks are visible. This is due to the paramagnetic relaxation enhancement (PRE) experienced

by the nuclei in the proximity of the metal. Addition of Mn^{2+} (SI Fig. S9A) results in the disappearance of several backbone amide crosspeaks ~7.7–8.7 ppm, however the peaks for the side amides of the two tryptophans (~10 ppm) appear only to show a slight decrease in intensity. The Co^{2+} bound form shows at least a few well dispersed backbone amide crosspeaks ~7.2–7.7 ppm and ~9.5 ppm (SI Fig. S9B). Moreover, the peaks for the side chain amides of the two tryptophans (W107 & W127, ~10 ppm), which are in or sequentially close to the C-X-X-C-G motifs (^{106}C-W-N-C-G^{110} & ^{128}C-P-Y-C-G^{132}) disappear of shift. Addition of Fe^{3+} ions (SI Fig. S9C) results in the disappearance or shift of a few backbone resonances. Although, the originally present peaks for the side chain amides of the two tryptophans (~10 ppm) appear only to show a slight decrease in intensity, clearly a new peak becomes visible ~10.5 ppm. Spectral differences between the different metal bound states may arise from differences in the affinity for the respective metals as well as differences regarding the paramagnetic properties of the coordinated metal.

Supplementary figure legends

Fig. S1: Superposition of the ^{1}H-^{15}N-HSQC spectra of His-PknG1-147 (black), His-PknG1-75 (green), and PknG74-147 (red). The RD was always present in the reduced, metal bound (Zn^{2+}) state. Note that the spectra of His-PknG1-75 (green), and PknG74-147 (red) are almost perfectly additive subspectra of the one of His-PknG1-147 (black). The assignments for His-PknG1-75, which corresponds to the natively disordered NORS region, are indicated by the one letter amino acid code and the sequence positions. The black rectangle contains side chain amide protons of glutamine and asparagine residues. Assignments for the RD can be found in Fig. 3A and SI Fig. S5. The assigned chemical shifts for all have been deposited at the BMRB (accession numbers 26028 for the His-PknG1-147 with the RD in reduced, metal bound state, 26027 for His-PknG1-75, 26030 and 26029 for PknG74-147 either in the reduced, metal bound or oxidized state, respectively) (6).

Fig. S2: More NMR data sensitive to the secondary structure content for the PknG N-terminal regions. **(A)** $^{3}J_{HNH\alpha}$ coupling constants of ^{15}N-His-PknG1-75 (NORS, blue) and ^{15}N-PknG74-147 (RD, red) plotted as a function of the residue sequence position. The data was derived from 3D HNHA spectra. The measured values were corrected by +11 % as suggested in the literature (7). Values below ~6-6.5 Hz are typically observed in α-helical regions, whereas values above ~8-8.5 Hz are characteristic for residues in β-sheets. Values in the

range ~6-8 Hz are typical for protein regions undergoing conformational exchange. The secondary structure elements for the RD presented above of the sequence were extracted from the crystal structure of PknG74-750 in complex with AX20017 (PDB-ID 2PZI, Fig. 1B, SI Fig. S6A) (4). **(B-C)** $^1H^\alpha$ secondary shifts for His-PknG1-75 (NORS, blue) and PknG74-147 (RD, reduced, metal bound, red) and $^{13}C^\alpha$ secondary shifts for the two segment protein His-PknG1-147 (NORS-RD, black) plotted as function of the residue sequence position (8). The $^{13}C^\alpha$ secondary shifts for the single region proteins are shown in Fig. 1C. Note, whereas $^1H^\alpha$ chemical shift values significantly lower than the random coil value indicate the presence of α-helical secondary structure elements and those significantly higher of β-sheets, the behavior is opposite for the $^{13}C^\alpha$ chemical shifts.

Fig. S3: (A) Further ^{15}N-relaxation data for ^{15}N-His-PknG1-75 (NORS, blue), ^{15}N-PknG74-147 (RD, reduced, metal bound, red), and ^{15}N-His-PknG1-147 (NORS-RD, reduced, metal bound, black). The top panel shows the ^{15}N-T_1 and the bottom the -T_2 values as a function of the residue sequence position. **(B)** {^1H}-^{15}N-NOE NMR data for PknG 74-147 (RD) in the oxidized, unstructured state. Positive and negative peaks of the reference spectrum are shown in black and blue, respectively, and those of the NOE spectrum in magenta and red, respectively. Red peaks on top of black ones indicate the presence of highly flexible regions with negative NOE values, whereas black peaks with no peak on top indicate less dynamic residues with NOE values around 0. Magenta peaks on top of black peaks indicate residues with positive NOE values typical for more structured regions. Note that the behavior for the side chain amide protons of the tryptophan peaks (^1H shift ~10 ppm) is opposite because they are spectrally folded. Assigned backbone amide crosspeaks that shifted or showed a change in signal intensity are labeled by the one-letter amino acid code and the sequence position (6).

Fig. S4: Comparison of the secondary structure content of the isolated PknG RD in solution compared to the three available crystal structures of constructs containing the RD and the kinase domain based on experimental and back calculated $^{13}C^\alpha$ secondary shifts. The experimental data for the Zn^{2+} form measured by NMR spectroscopy is shown in the top panel (red bars) (8). $^{13}C^\alpha$ secondary shifts based on the published crystal structures were calculated using the program SPARTA+ (9). The second panel (yellow bars) shows the data for the Cd^{2+} bound form of the RD in the structure of PknG74-750 (RD-KD-TPRD) in complex with the inhibitor AX20017 (PDB-ID 2PZI) (4). Those for the Zn^{2+} bound forms of

the RD in the structures of PknG74-405 (RD-KD) in complex with either an ATP analogue (ATP-γS, blue bars, PDB-ID 4Y12) or ADP (green bars, PDB-ID 4Y0X) (5) are shown in the third and fourth panel, respectively. Chemical shift values significantly higher than the random coil value indicate the presence of α-helical structure elements and those significantly lower of β-sheets. The above indicated secondary structure elements were extracted from the respective crystal structures. See also Fig. 1C and SI Fig. S2.

Fig. S5: NMR data regarding the comparison of the structures of the metal bound rubredoxin-like domain in solution in the isolated form and in the presence of the following kinase domain. The picture shows a superposition of the ^{1}H-^{15}N-HSQC spectra of ^{15}N-D-His-PknG74-420 (RD-KD, black) and ^{15}N-PknG74-147 (RD, red). The RD was always present in the reduced, Zn^{2+} bound state. The assignments for PknG74-147 are indicated by the one-letter amino acid code and the sequence positions (6). The presence of the His-tag at the N-terminus and of the kinase C-terminus of the RD in His-PknG74-420 is expected to result in different chemical shifts for the RD N- and C-terminal regions compared to the tag free isolated RD. Due to the significantly larger size and the fact that His-PknG74-420 is deuterated some backbone resonances may not be visible because they are broadened beyond detection or because the back exchange of undetectable deuterons to detectable protons did not occur equally well for less solvent accessible backbone amides.

Fig. S6: More data regarding the comparison of the structures of the isolated metal bound rubredoxin-like domain in solution and in the crystal structures of PknG fragments containing additionally the kinase domain. **(A)** Ribbon representations of the RD-KD regions of the three crystal structures. The RD region is colored according to the B-factors (red - small B factors to blue - high B factors). Some residues that are dynamic had no coordinates in the respective pdb file of the crystal structure. The kinase domain is represented in yellow. The B-factors were plotted using the software Pymol with a minimum value of 20 (red) and a maximum value of 50 (blue). **(B)** In the amino acid sequence of the rubredoxin-like motif (RD) shown at the top, residues that were used for the determination of the alignment tensor for the back calculation of the RDCs from the crystal structures are highlighted in red. The spatial position of the respective amino acids in the structure of the RD (based on PDB-ID 4Y0X) is indicated in the ribbon representation. Plots of the linear correlation between experimental RDCs (partial alignment with 17 mg/ml phages PF1) and those back calculated

based on the three published crystal structures using the software PALES (10) are shown below and to the right. The data for the Cd^{2+} bound RD of PknG74-750 in complex with the inhibitor AX20017 (PDB-ID 2PZI) (4) are represented as green symbols and those for the Zn^{2+} bound RD of PknG74-405 in complex with an ATP analogue (PDB-ID 4Y12) (5) as red symbols and in complex with ADP (PD-ID 4Y0X) (5) as blue symbols. The bottom two plots compare the back calculated RDCs of the three crystal structures (left: PknG74-405 in complex with ADP or an ATP analogue versus PknG74-750 in complex with an inhibitor, right: PknG74-405 in complex with ADP versus the complex with an ATP analogue). Whereas the structure of the RD in the two nucleotide bound structures is highly similar (right plot), it differs somewhat between the inhibitor bound state of PknG74-750 and the nucleotide bound states of PknG74-405 (left plot).

Fig. S7: More kinase assay data for different PknG constructs and redox conditions. **(A) & (C)** SDS-PAGE analysis of the *in vitro* kinase assays using radio labeled ATP and 15% polyacrylamide gels. The gels were analyzed with a phosphorimager. **(A)** shows the gel picture using as kinases His-PknG wild type (wt, blue), His-PknG74-750 (green), or His-PknG74-420 (red) and as substrate His-PknG1-147 with the RD in the oxidized, metal free or reduced, metal bound form. **(B)** Comparison of kinase activities of PknG wild type and of an N-terminal truncated version (PknG74-750) after treatment of their RD metal binding motif with different redox conditions. Kinase activities for the wild type enzyme are represented including the auto phosphorylation in trans of its own N-terminus (see upper band on gel picture in (C)). **(C)** shows the respective SDS PAGE analysis using as kinases His-PknG wild type (wt, blue & red) or His-PknG74-750 (green & black) with the RD either in the reduced, metal bound or oxidized, metal free state and as substrate His-PknG1-147 also with the RD either in the oxidized, metal free or reduced, metal bound form. Please, see also the labeling in each figure. **(D)** NMR monitoring of *in vitro* ^{15}N-His-PknG1-75 (NORS) phosphorylation by His-PknG74-420 (RD-KD) based on the superposition of ^{1}H-^{15}N-HSQC spectra of unphosphorylated ^{15}N-His-PknG1-75 (black), after kinase treatment for 3 h at 298 K (red spectra) and further overnight at 310 K (green). A * indicates a peak that appears newly after phosphorylation. Assigned backbone amide crosspeaks that shifted or showed a change in signal intensity are labeled by the one-letter amino acid code and the sequence position (see also SI Fig. S1) (6). The additional label –sc indicates side chain amide protons.

Fig. S8: More NMR data regarding the effect of oxidization of the two RD C-X-X-C-G motifs on the spectral appearance. **(A)** Superposition of the ^{1}H-^{15}N-HSQC spectra of PknG74-147 (RD) in the reduced, metal bound folded state (black) and the oxidized, unfolded state (red). **(B)** Superposition of the ^{1}H-^{15}N-HSQC spectra of the oxidized, unfolded states of His-PknG1-147 (NORS-RD, black) and PknG74-147 (RD, red). **(C)** Superposition of the ^{1}H-^{15}N-HSQC spectra of the oxidized, unfolded state of His-PknG1-147 (black) and of His-PknG1-147-4C/S mutant (red).

Fig. S9: More NMR data regarding the coordination of different metal ions by the two RD C-X-X-C-G motifs. **(A-C)** Superposition of the ^{1}H-^{15}N-HSQC spectra of ^{15}N-His-PknG1-147 (RD) in the oxidized, unfolded state (black) and reduced, metal bound state (red), the metal ion being either Mn^{2+}, Co^{2+}, or Fe^{3+}, respectively. The reducing agent was always TCEP and the counterion Cl$^-$.

Fig. S10: (A) Last snapshots from six independent 250 ns MD simulations of PknG74-420 with the cysteines of the two C-X-X-C-G motifs of the RD in the reduced state and coordinating Fe^{2+} (Fe^{2+} bound) or in the oxidized state with disulfide bridges between C106 and C109, and C128 and C131 (oxidized). The core RD region (residues ~100-140) is shown in red. Protein atoms within 5 Å of ATP and Mg^{2+} forming the binding cavity are represented in space filling mode in cyan. The ATP binding site is on average more accessible and the RD domain is less folded in the oxidized simulations as compared to the Fe^{2+} bound state. **(B)** Distribution of Loop 2 extension (residue 292-297, shown in cyan in Fig. 5B) calculated form three independent 250 ns MD simulations of the Fe^{2+} bound (in blue) and oxidized (in red) states.

Fig. S11: (A) The volume of the ATP binding cavity, and **(B)** solvent accessible surface area (SASA) of residues interacting with ATP (I87, A92, I157, A158, I165, V179, K181, M232, Y234, G236, M283, I292, D293)(4) in the Fe^{2+} bound (left) and oxidized (right) states.
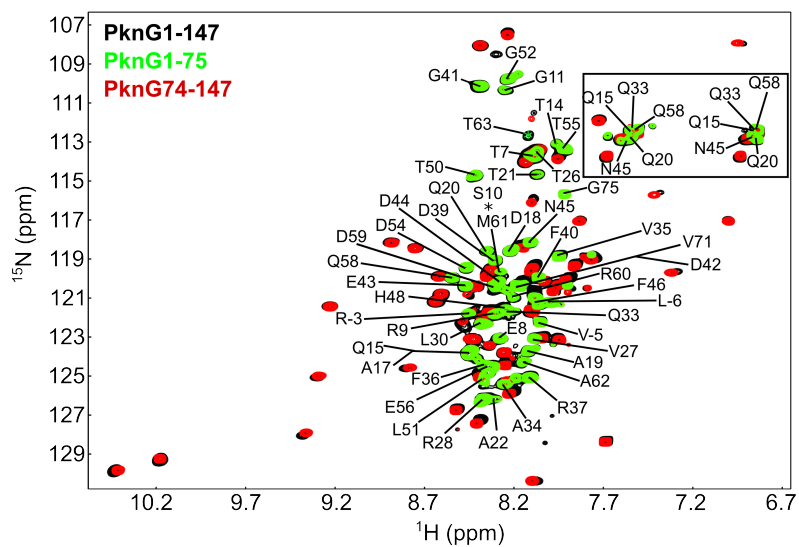
Fig. S12: Data from 250 ns MD simulations of PknG74-420 with the RD in the reduced metal (Fe^{2+}) bound state or the oxidized, metal free state for the characterization of the influence of RD oxidation on the conformation of the RD as well as the kinase. **(A)** Stick representation of residues surrounding the ATP-Mg^{2+} binding pocket. Only one coordination bond remains stable between Mg^{2+} and E280 during the MD trajectory in the Fe^{2+} bound state
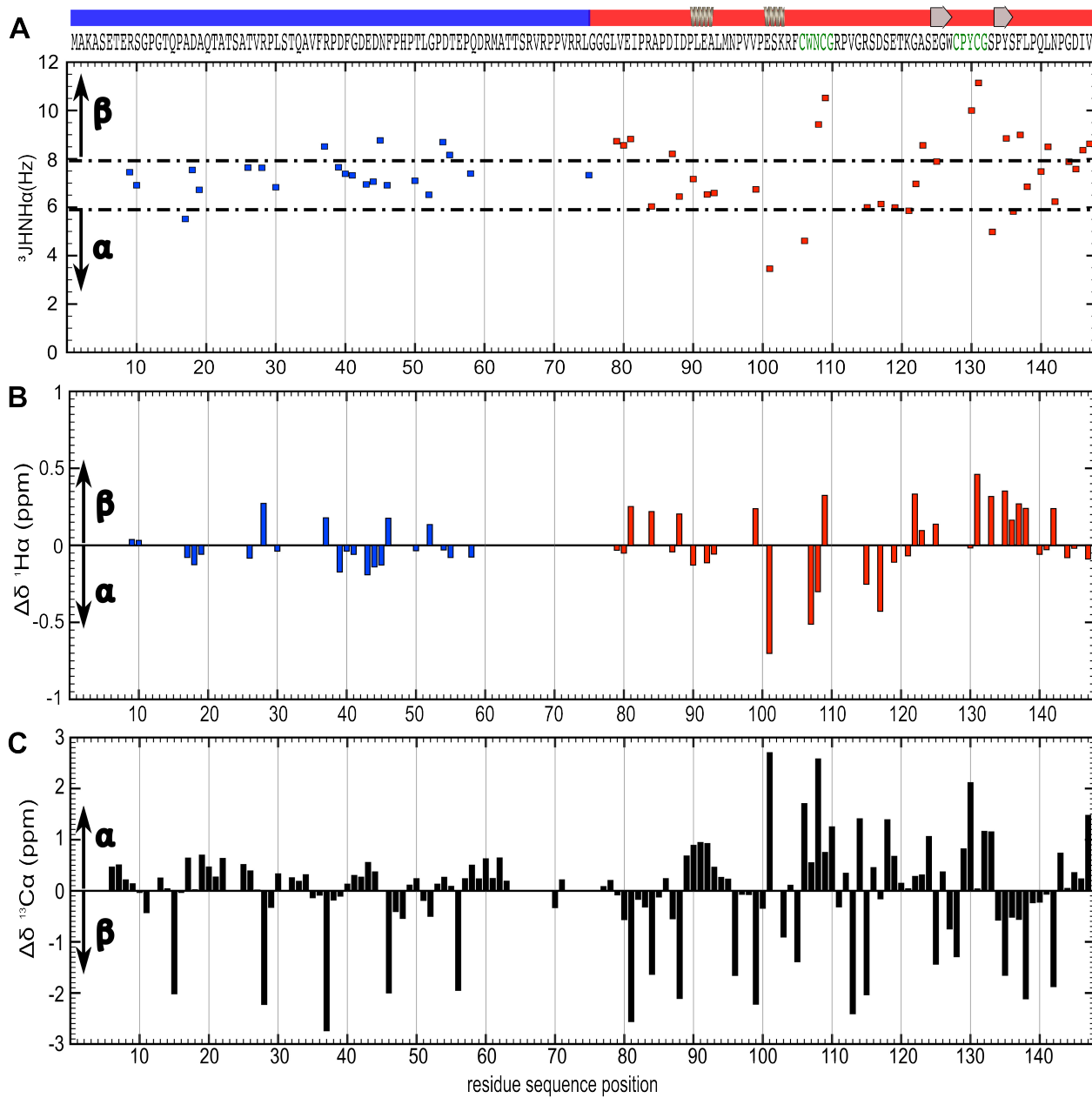
(left). In the oxidized state, Mg^{2+} forms two coordination bound with surrounding residues E280 and D293 (right). **(B)** The distance between Mg^{2+} and D293 as a function of the simulation time for the oxidized and Fe^{2+} bound state. **(C)** The distance between the functionally important residue D276 of the catalytic loop and K278 is larger in the oxidized compared to the Fe^{2+} bound state, which may result in an increased catalytic activity of the former state. In the three independent 250 ns simulations, the D276-K278 ion pair is open (> 4 Å) in 6%, 23%, and 47%, respectively, of the simulation time for the Fe^{2+} bound state, and 4%, 37%, 80%, respectively, of the simulation time in the oxidized state.
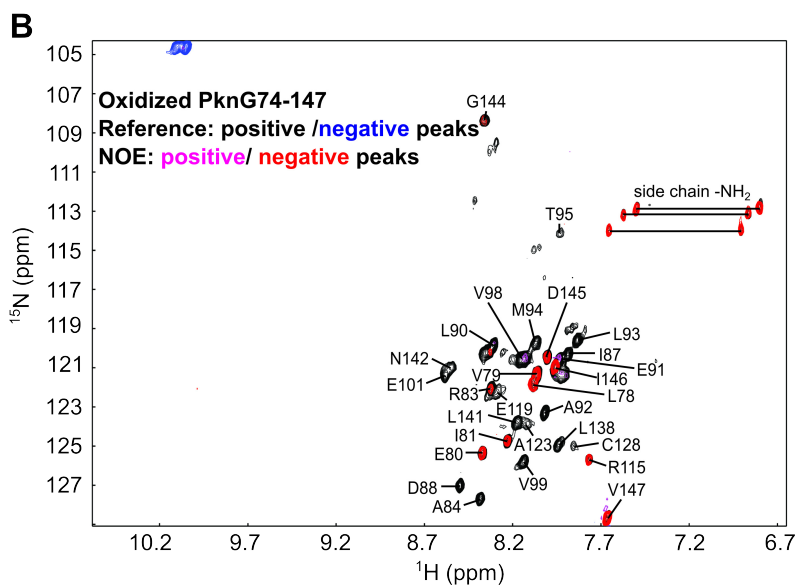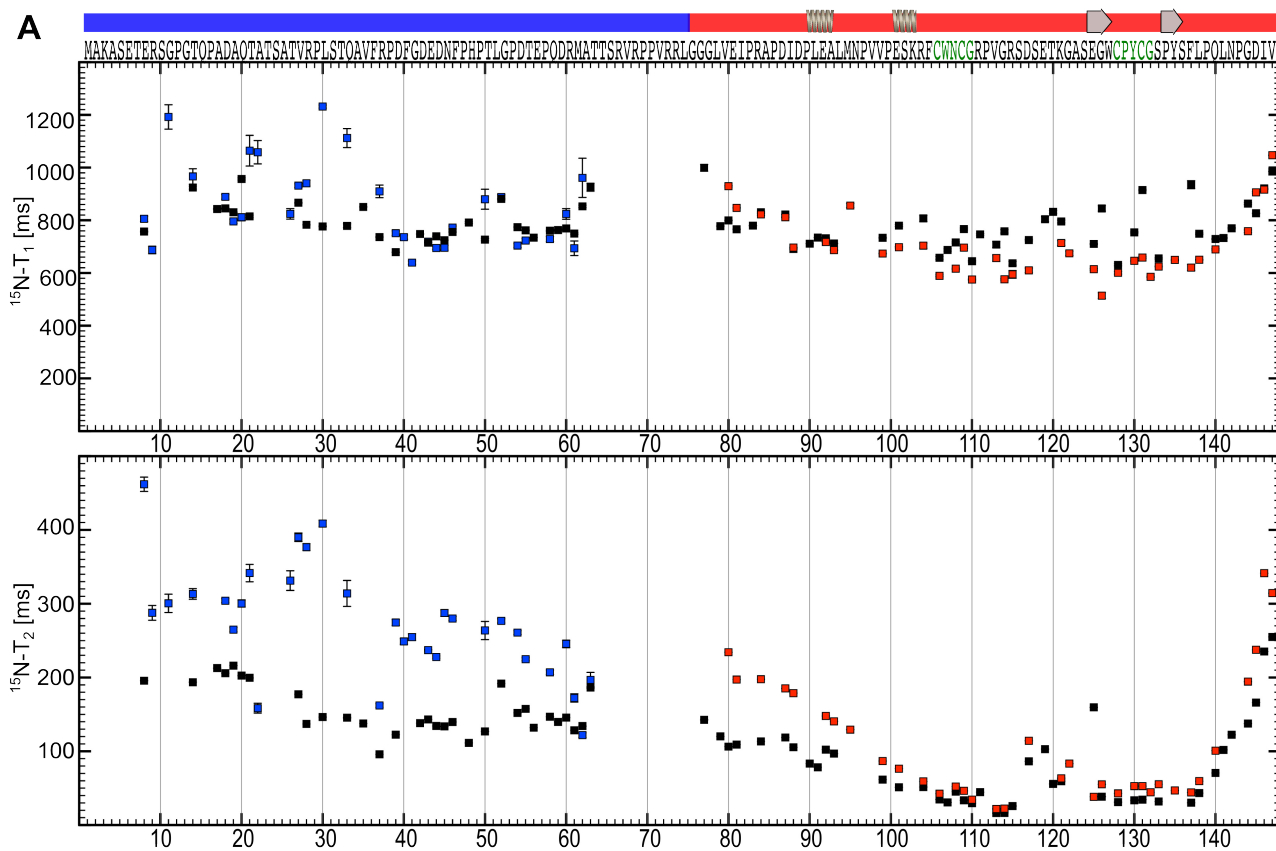
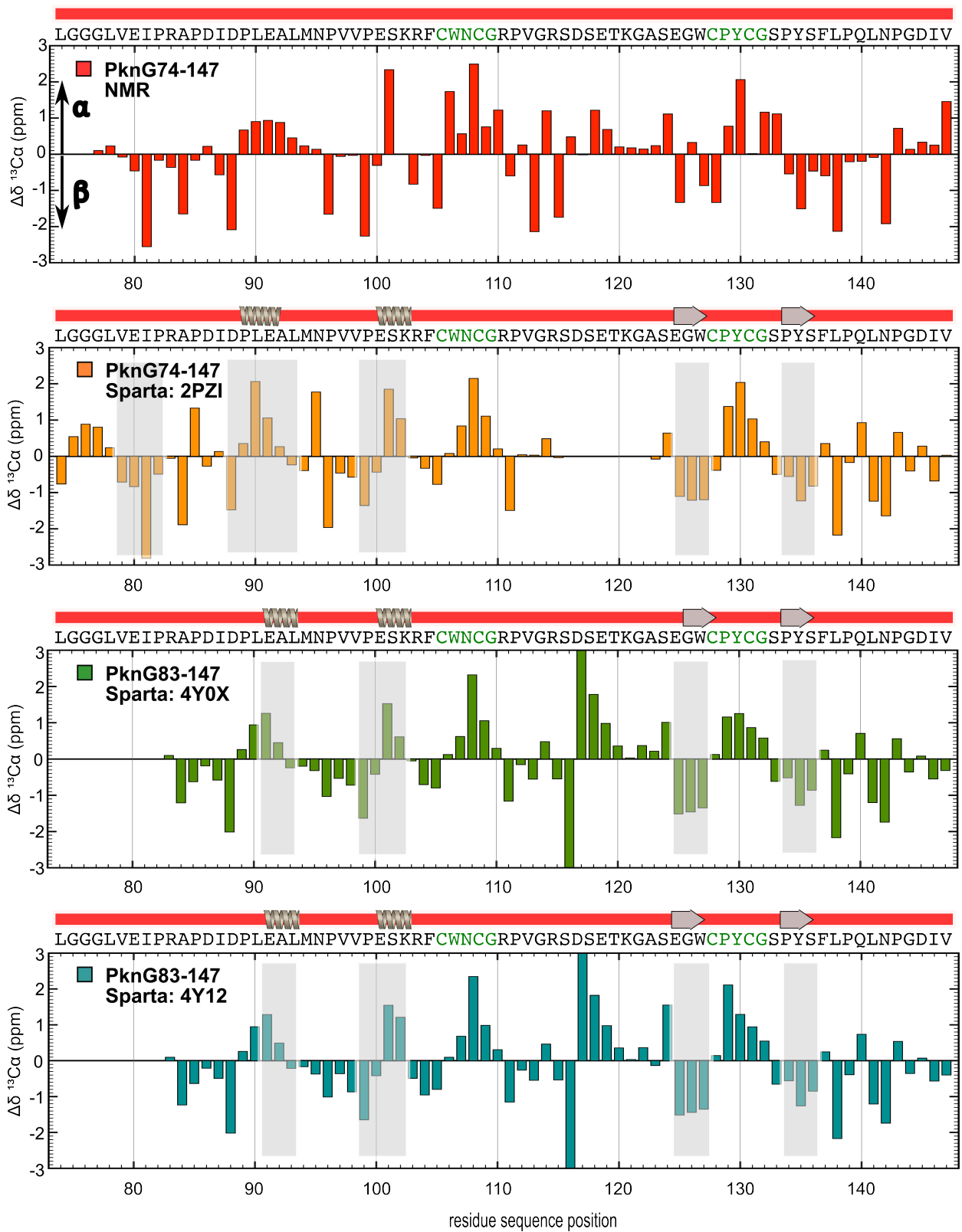**Supplementary References**

1.   Walsh, J. D., Meier, K., Ishima, R., and Gronenborn, A. M. (2010) NMR studies on domain diffusion and alignment in modular GB1 repeats. *Biophys J* **99**, 2636-2646
2.   Dames, S. A., Junemann, A., Sass, H. J., Schonichen, A., Stopschinski, B. E., Grzesiek, S., Faix, J., and Geyer, M. (2011) Structure, dynamics, lipid binding, and physiological relevance of the putative GTPase-binding domain of Dictyostelium formin C. *The Journal of biological chemistry* **286**, 36907-36920
3.   Sieker, L. C., Stenkamp, R. E., and LeGall, J. (1994) Rubredoxin in crystalline state. *Methods in enzymology* **243**, 203-216
4.   Scherr, N., Honnappa, S., Kunz, G., Mueller, P., Jayachandran, R., Winkler, F., Pieters, J., and Steinmetz, M. O. (2007) Structural basis for the specific inhibition of protein kinase G, a virulence factor of Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 12151-12156
5.   Lisa, M. N., Gil, M., Andre-Leroux, G., Barilone, N., Duran, R., Biondi, R. M., and Alzari, P. M. (2015) Molecular Basis of the Activity and the Regulation of the Eukaryotic-like S/T Protein Kinase PknG from Mycobacterium tuberculosis. *Structure* **23**, 1039-1048
6.   Wittwer, M., and Dames, S. A. (2016) Chemical shift assignment of the intrinsically disordered N-terminus and the rubredoxin domain in the folded metal bound and unfolded oxidized state of mycobacterial protein kinase G. *Biomol NMR Assign* **10**, 401-406
7.   Vuister, G. W., and Bax, A. (1993) Quantitative J correlation: a new approach for measuring homonuclear three-bond J(HNH.alpha.) coupling constants in 15N-enriched proteins. *Journal of the American Chemical Society* **115**, 7772-7777
8.   Wishart, D. S., and Sykes, B. D. (1994) The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. *Journal of biomolecular NMR* **4**, 171-180
9.   Shen, Y., and Bax, A. (2010) SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. *Journal of biomolecular NMR* **48**, 13-22
10.   Zweckstetter, M. (2008) NMR: prediction of molecular alignment from structure using the PALES software. *Nature protocols* **3**, 679-690

A

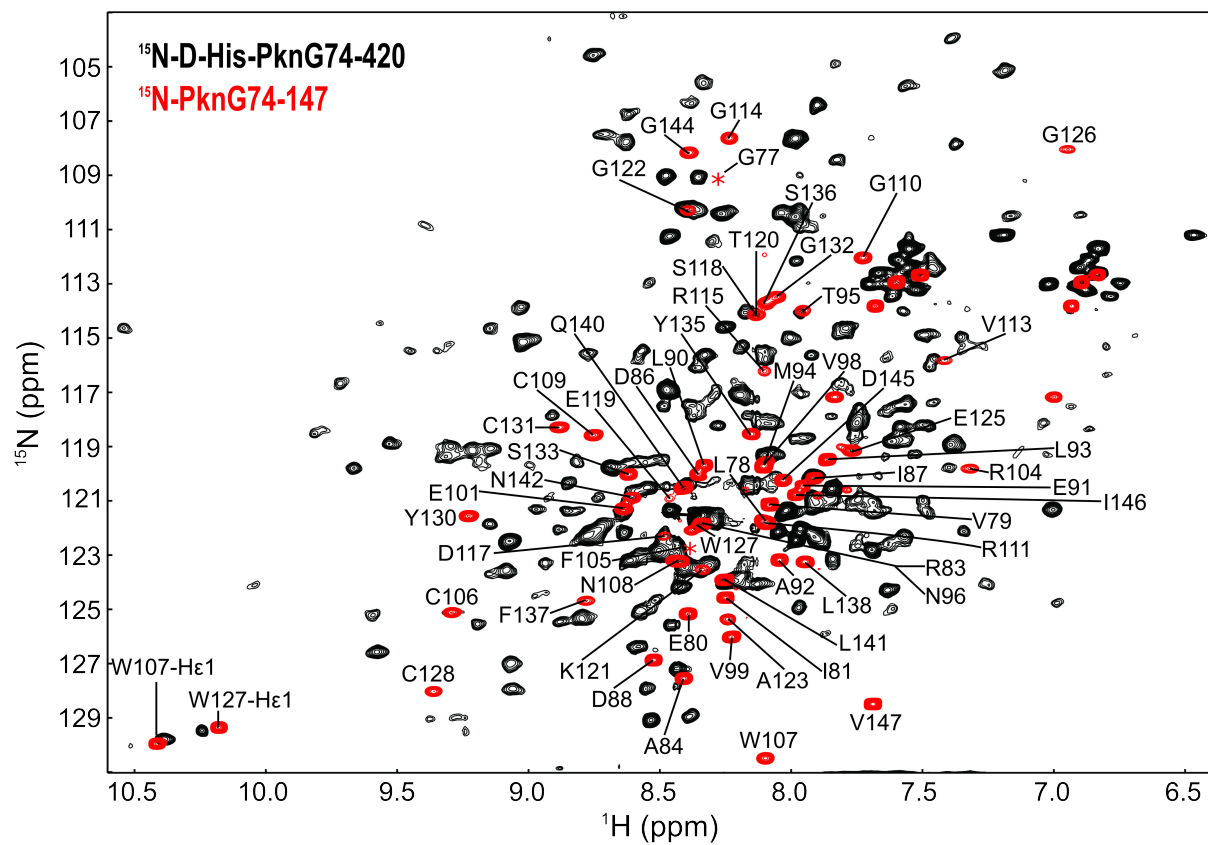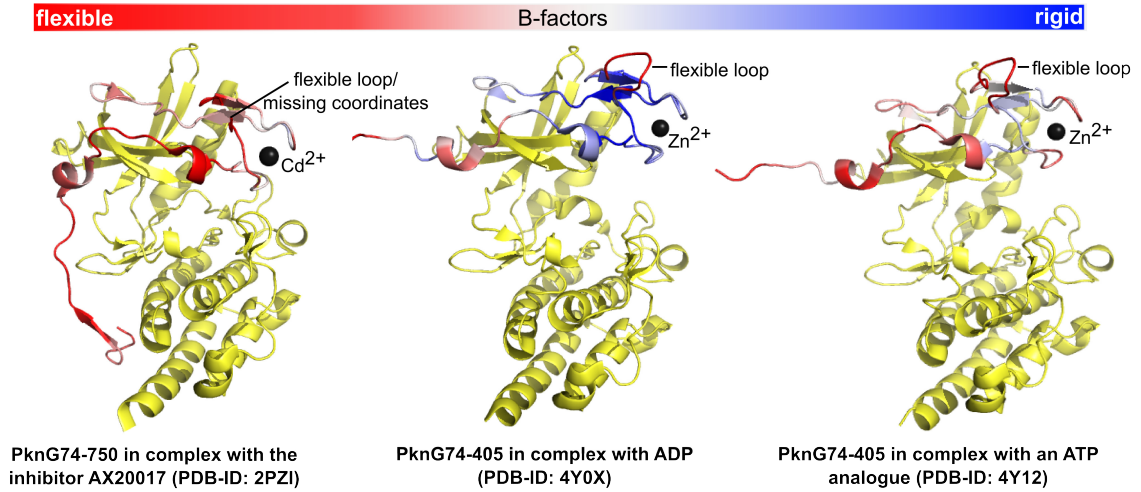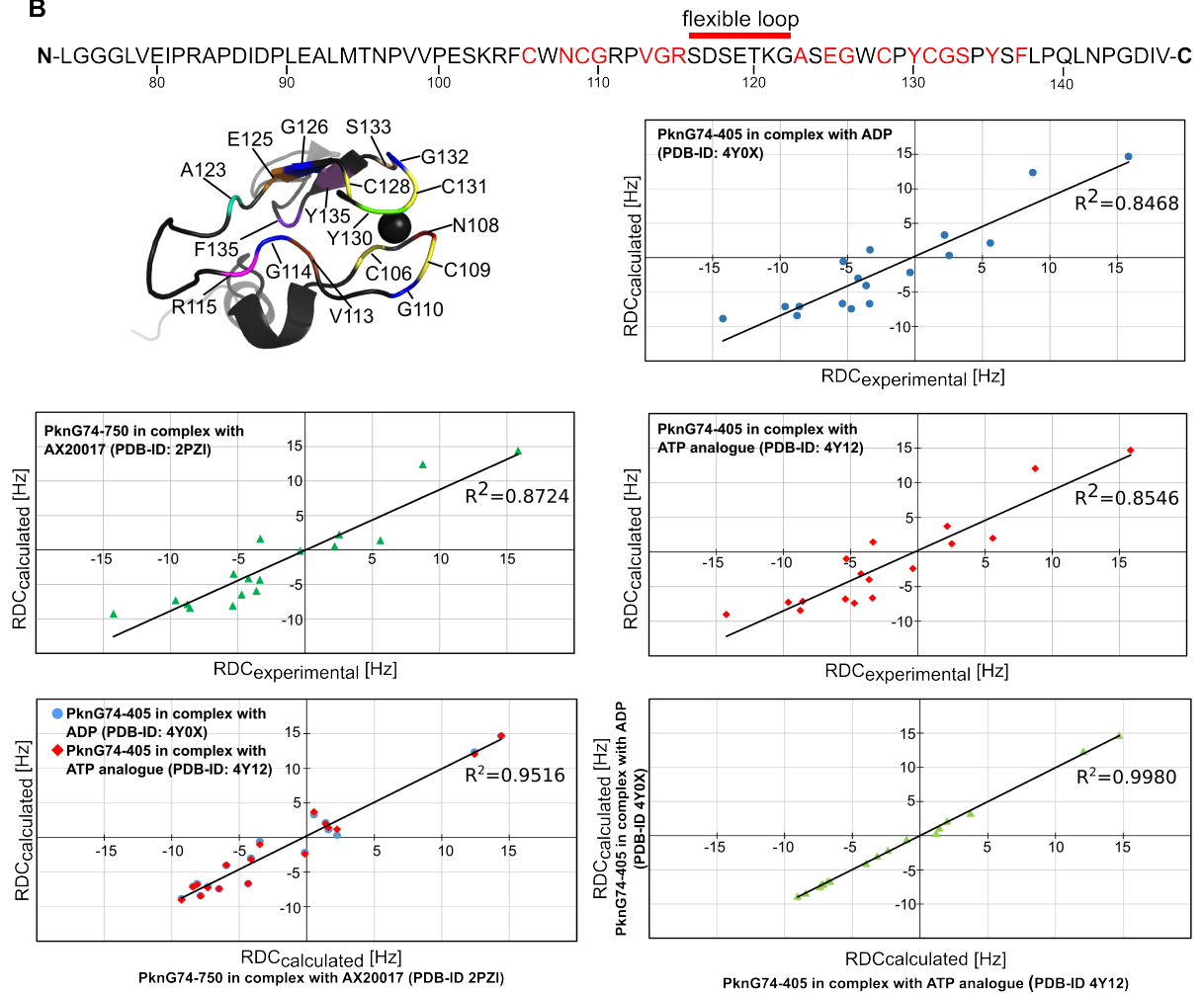MAKASETERSGPGTQPADAQTATSATVRPLSTQAVFRPDFGDEDNFPHPTLGPDTEPQDRMATTSRVRPPVRRLGGGLVEIPRAPDIDPLEALMNPVVPESKRFCWNCGRPVGRSDSETKGASEGWCPYCGSPYSFLPQLNPGDIV

B

C

A

PknG74-750 in complex with the
inhibitor AX20017 (PDB-ID: 2PZI)

PknG74-405 in complex with ADP
(PDB-ID: 4Y0X)

PknG74-405 in complex with an ATP
analogue (PDB-ID: 4Y12)

B

flexible loop

N-LGGGLVEIPRAPDIDPLEALMTNPVVPESKRF CWNCG RP VGR SDSETKGA SE GW C PYCGS PYS F LPQLNPGDIV-**C**

80   90   100   110   120   130   140

PknG74-405 in complex with ADP
(PDB-ID: 4Y0X)

$R^2=0.8468$

PknG74-750 in complex with
AX20017 (PDB-ID: 2PZI)

$R^2=0.8724$

PknG74-405 in complex with
ATP analogue (PDB-ID: 4Y12)

$R^2=0.8546$

PknG74-405 in complex with
ADP (PDB-ID: 4Y0X)
PknG74-405 in complex with
ATP analogue (PDB-ID: 4Y12)

$R^2=0.9516$

PknG74-750 in complex with AX20017 (PDB-ID 2PZI)

$R^2=0.9980$

PknG74-405 in complex with ATP analogue (PDB-ID 4Y12)

A

Kinase:
PknG wt (NORS-RD/Zn-KD-TPRD)
PknG74-750 (RD/Zn-KD-TPRD)
PknG74-420 (RD/Zn-KD)

-> autophosphorylation

-> substrate phosphorylation

PknG1-147 (NORS-RD/ox)
PknG1-147 (NORS-RD/Zn)
PknG1-147 (NORS-RD/ox)
PknG1-147 (NORS-RD/Zn)
PknG1-147 (NORS-RD/ox)
PknG1-147 (NORS-RD/Zn)

B

Kinase:
PknG74-750 (RD/ox-KD-TPRD)
PknG wt (NORS-RD/ox-KD-TPRD)
PknG74-750 (RD/Zn-KD-TPRD)
PknG wt (NORS-RD/Zn-KD-TPRD)

C

Kinase:
PknG wt (NORS-RD/Zn-KD-TPRD)
PknG wt (NORS-RD/ox-KD-TPRD)
PknG74-750 (RD/ox-KD-TPRD)
PknG74-750 (RD/Zn-KD-TPRD)

-> autophosphorylation

-> substrate phosphorylation

PknG1-147 (NORS-RD/ox)
PknG1-147 (NORS-RD/Zn)
PknG1-147 (NORS-RD/Zn)
PknG1-147 (NORS-RD/ox)
PknG1-147 (NORS-RD/Zn)
PknG1-147 (NORS-RD/ox)
PknG1-147 (NORS-RD/ox)
PknG1-147 (NORS-RD/Zn)

D

Kinase:
PknG74-420 (RD/Zn-KD)
Substrate:
^{15}N-His-PknG1-75 (NORS)
0 h
3 h
overnight

A

^{15}N-PknG74-147
reduced / metal bound
^{15}N-^{13}C-PknG74-147
oxidized / metal free

B

^{15}N-His-PknG1-147
oxidized / metal free
^{15}N-^{13}C-PknG74-147
oxidized / metal free

C

^{15}N-His-PknG1-147
oxidized / metal free
^{15}N-His-PknG1-147-4C/4S

A

Fe²⁺ bound (1)   Fe²⁺ bound (2)   Fe²⁺ bound (3)
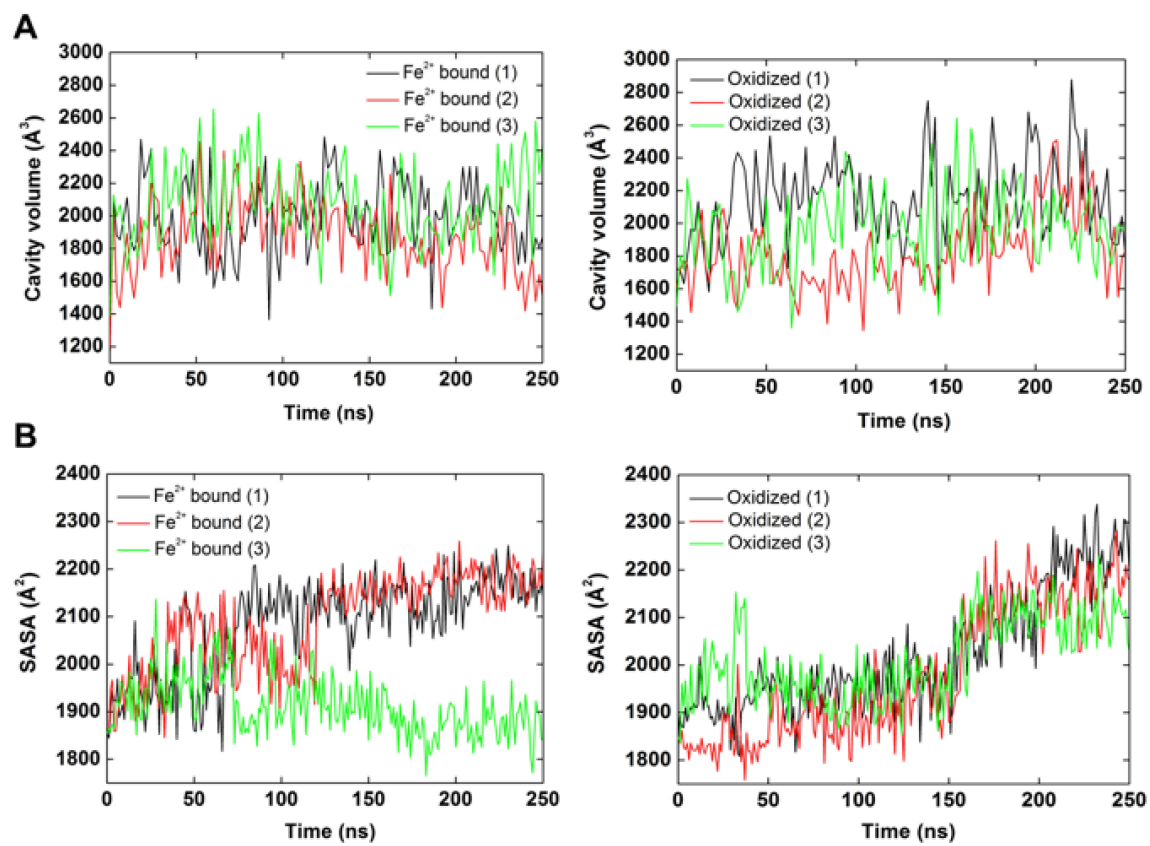
Oxidized (1)   Oxidized (2)   Oxidized (3)

B

A

B

C
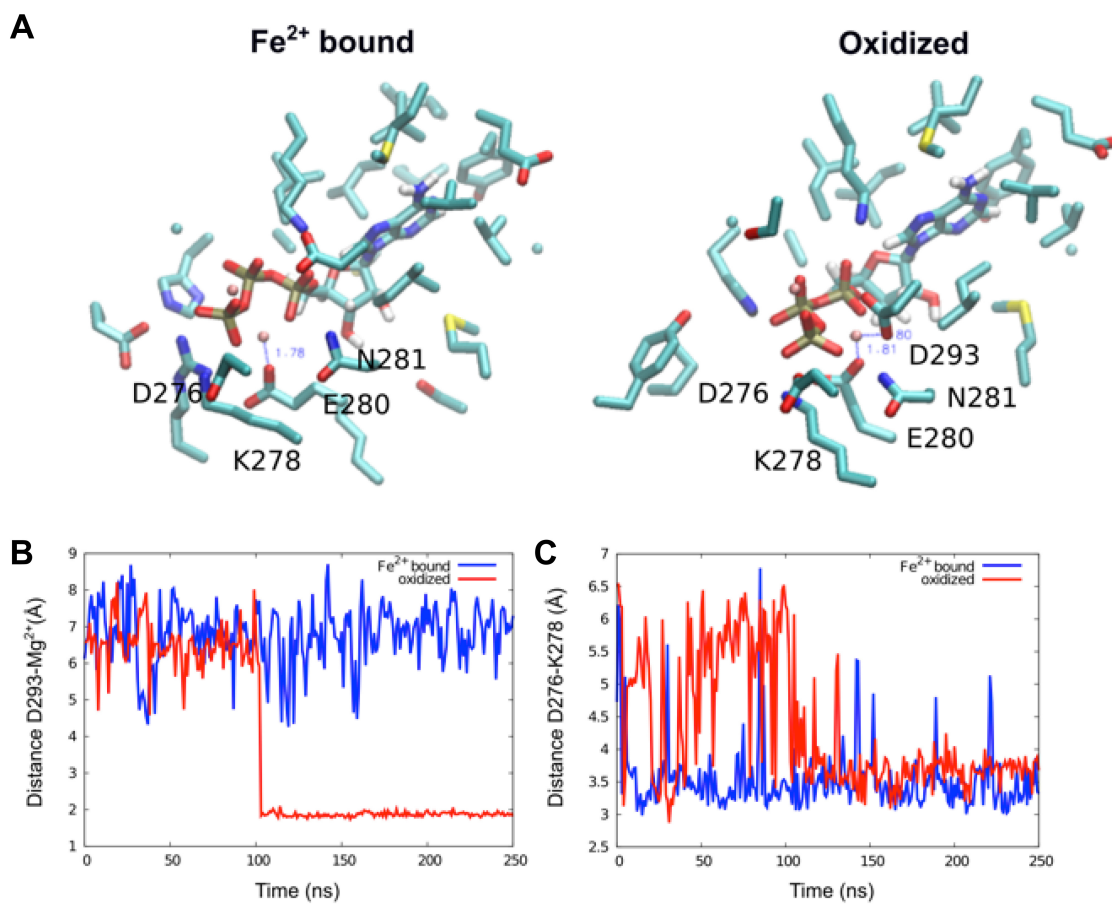
Acknowledgements - Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die durch ihre fachliche und persönliche Unterstützung zum Gelingen dieser Arbeit beigetragen haben.

Zunächst gilt mein herzlicher Dank PD Dr. Sonja A. Dames, die mir die Bearbeitung dieses interessanten Themas ermöglicht hat und mir dabei stets mit ihrer Expertise zur Seite stand. Weiterhin danke ich ihr für ihre unermüdliche, professionelle Einführung in die NMR und für die vielen konstruktiven, wissenschaftlichen Gespräche, die eine Grundlage bei der Realisierung dieser Arbeit waren.

Prof. Dr. Michael Sattler danke ich für seine Bereitschaft zur Übernahme des Zweitgutachtens und Prof. Dr. Bernd Reif für die Übernahme des Prüfungsvorsitzes. Darüber hinaus möchte ich mich bei Ihnen für die Unterstützung sowie für die Möglichkeit bedanken, Ihre Labore zu nutzen und an Ihren Seminaren und Ausflügen teilnehmen zu können.

Prof. Dr. Ville R. I. Kaila und Li Quo danke ich für die Durchführung der MD Simulationen sowie für ihren wissenschaftlichen Beitrag, welcher zum Gelingen der dritten Publikation beigetragen hat. Edgar Boczek danke ich für seine Hilfe während der Durchführung der Kinaseaktivitätsassays.

Den Mitarbeitern der Dames-, Reif- und Sattlergruppe danke ich für die schöne und konstruktive Zeit im Labor sowie während der gemeinsamen Seminare und Ausflüge. Mein besonderer Dank gilt dabei Sebastian, Lisa, Elke, Diana und Maristella für die vielen lustigen Stunden, die wir zusammen verbracht haben.

Meiner Familie und meiner Frau danke ich von ganzem Herzen dafür, dass sie mich bei allen meinen Entscheidungen unterstützt und mir geholfen haben meine Träume und Wünsche zu erfüllen. Auch in schwierigen Zeiten standen sie mir mit Rat und Tat zur Seite und hatten immer ein offenes Ohr für meine Probleme. Christiane, vielen Dank für dein Verständnis, deine Geduld und deine Hilfe. Ich bin so froh, euch an meiner Seite zu haben.

Ein ganz besonderer Dank gilt auch meinen Großeltern sowie Maria und Armin, die mir mit Ihren aufmunternden Worten geholfen haben auch in schwereren Zeiten das Ziel nicht aus den Augen zu verlieren. Ich werde euch für immer in meinem Herzen tragen.