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)** $^{1}H^{\alpha}$ 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 $^{1}H^{\alpha}$ 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)** $\{^{1}H\}$-^{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 (^{1}H 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.

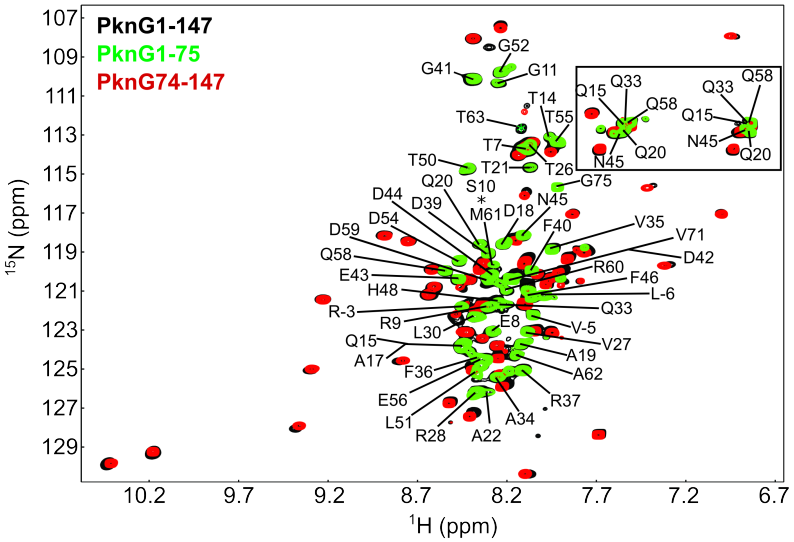
Fig. S12: Data from 250 ns MD simulations of PknG74-420 with the RD in the reduced metal (Fe^{2+}) bound state or the oxidized, metal free state for the characterization of the influence of RD oxidation on the conformation of the RD as well as the kinase. **(A)** Stick representation of residues surrounding 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.

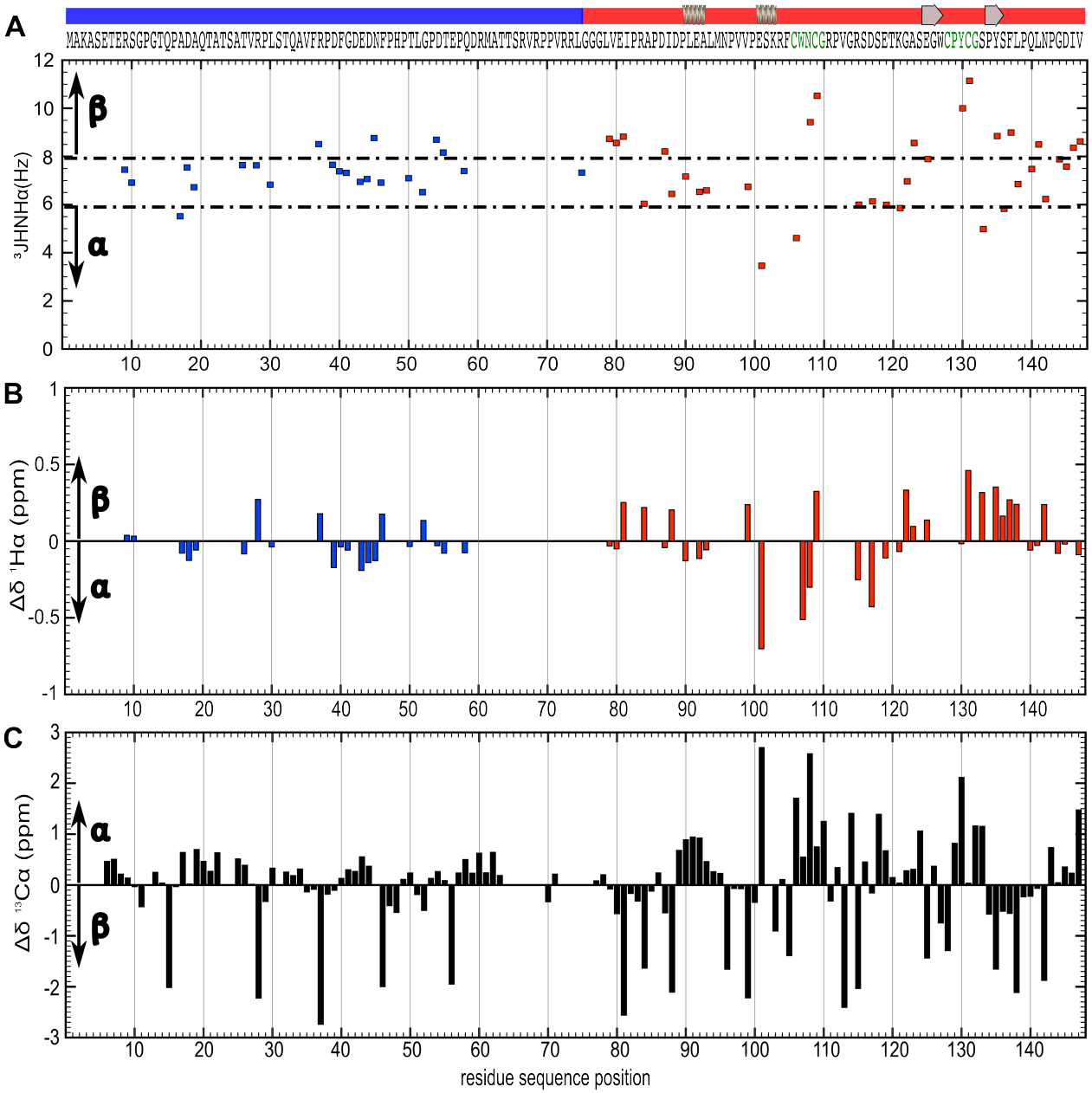
Supplementary References

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

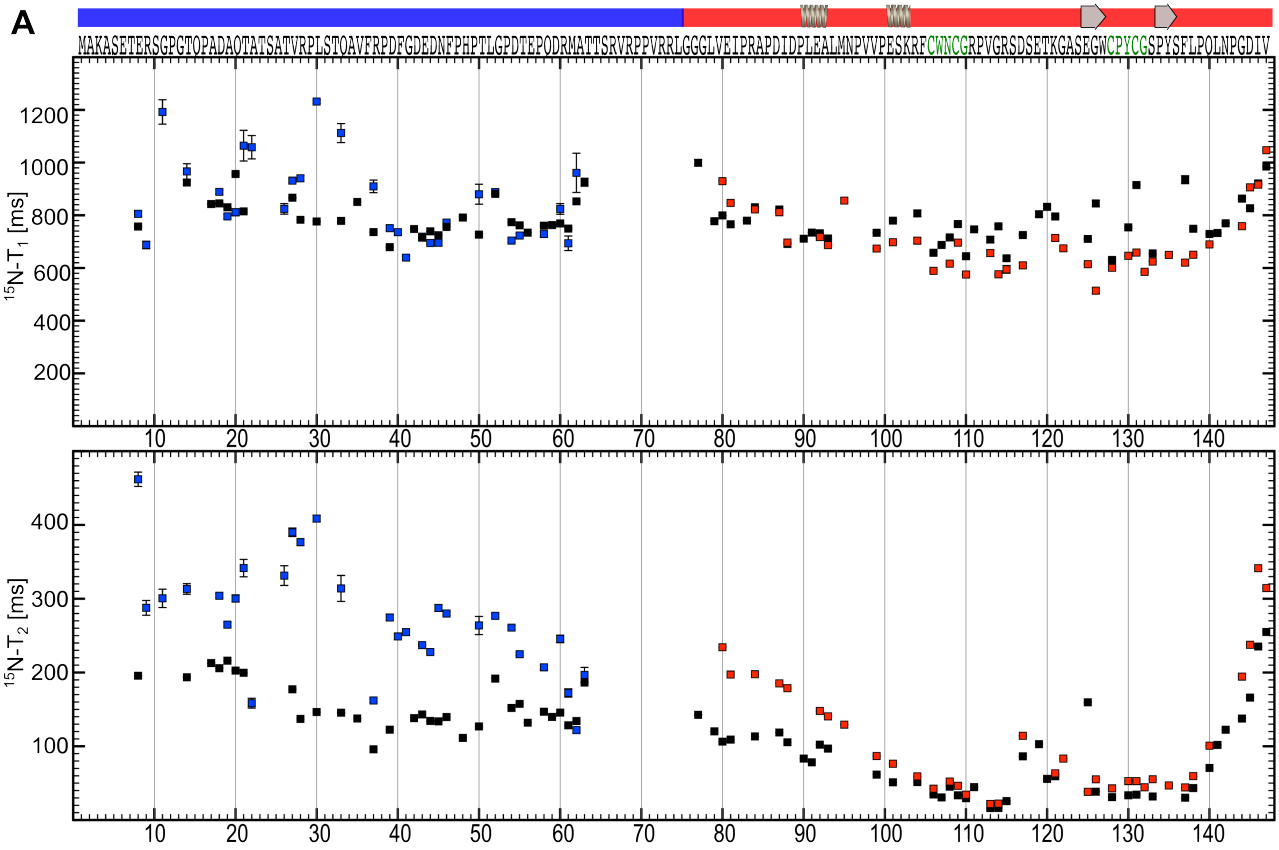
Wittwer et al.
SI Fig. S1

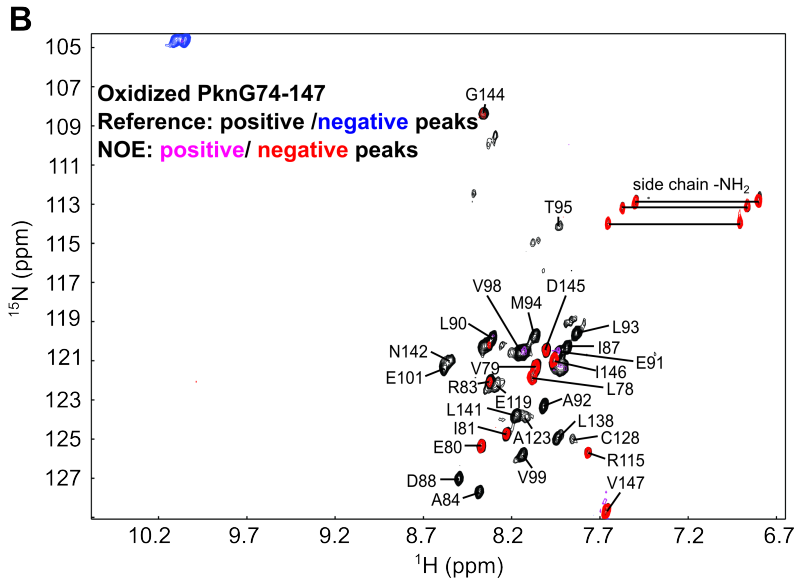


Wittwer et al.
SI Fig. S2

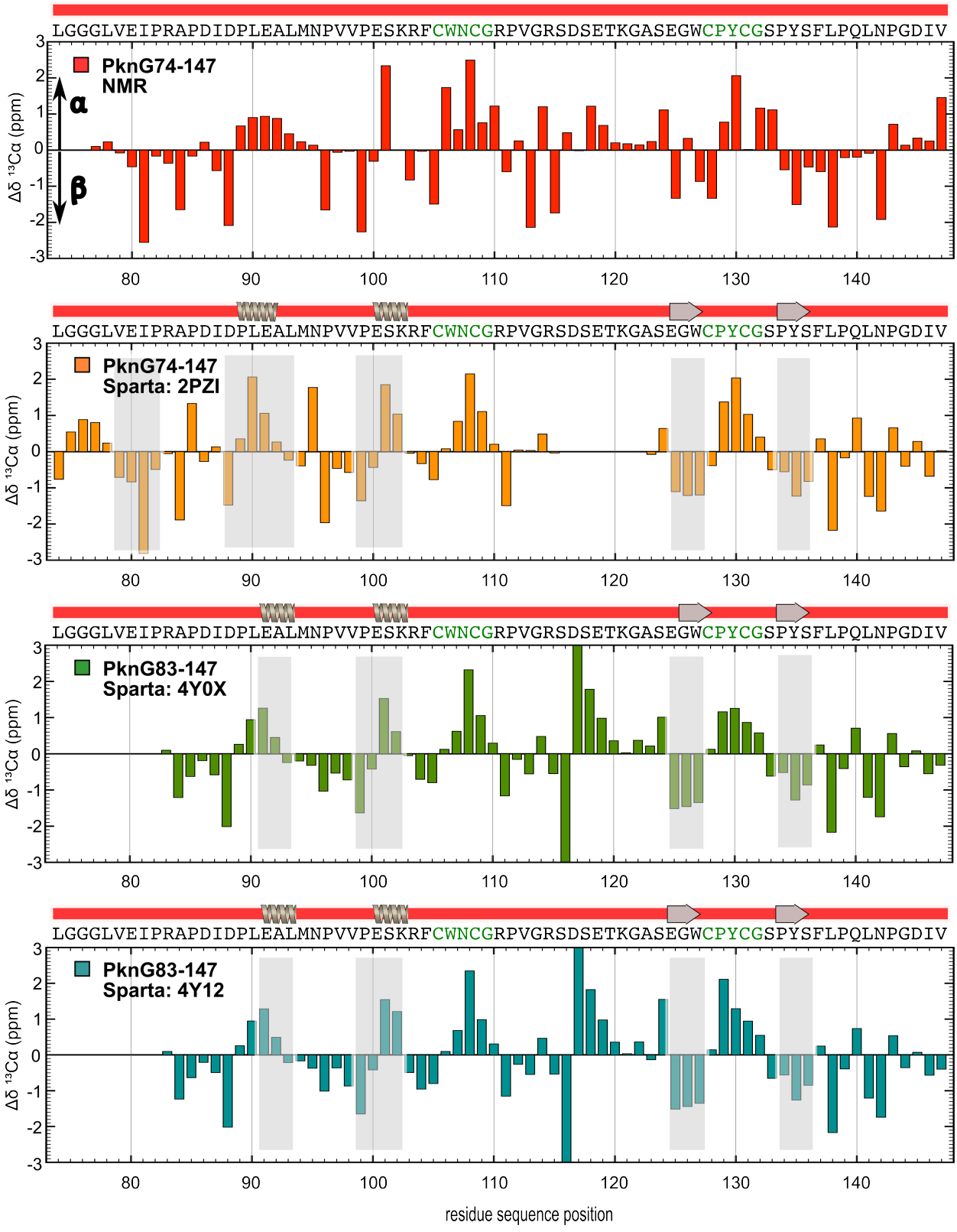


Wittwer et al.
SI Fig. S3

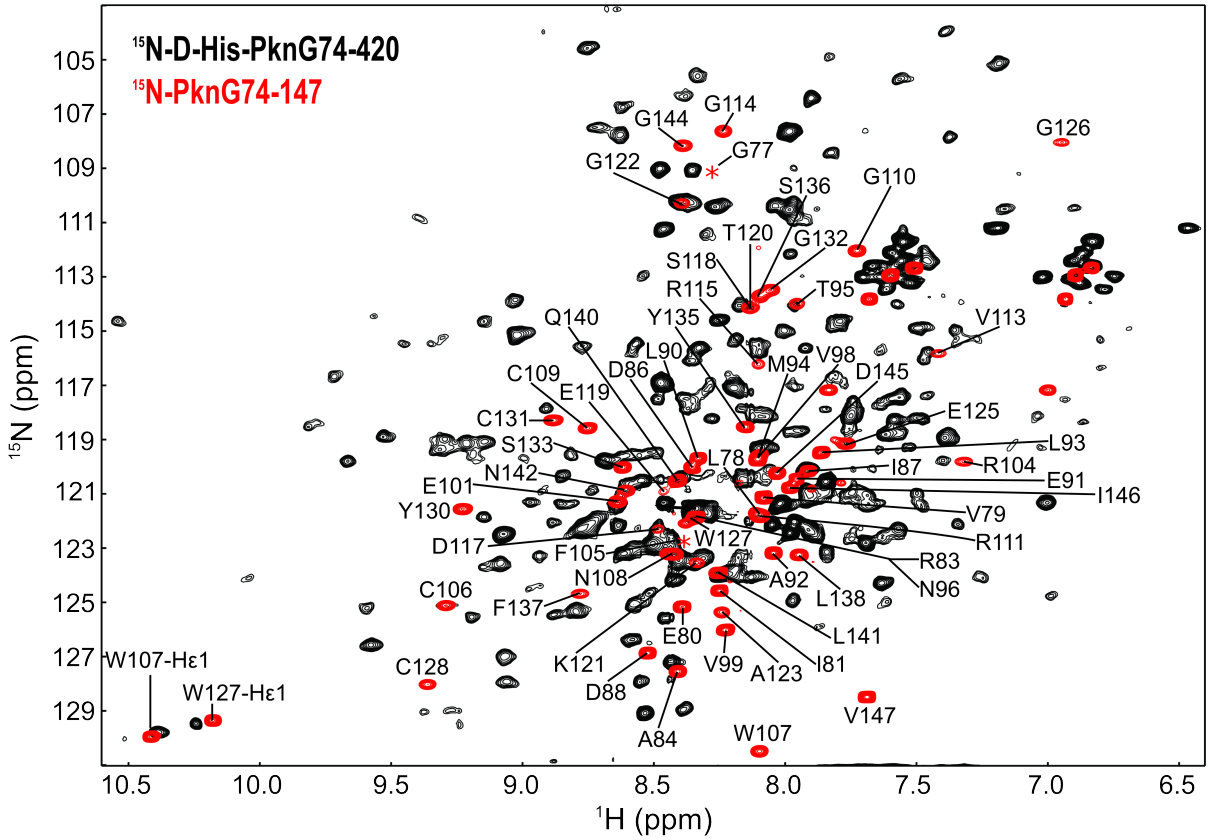




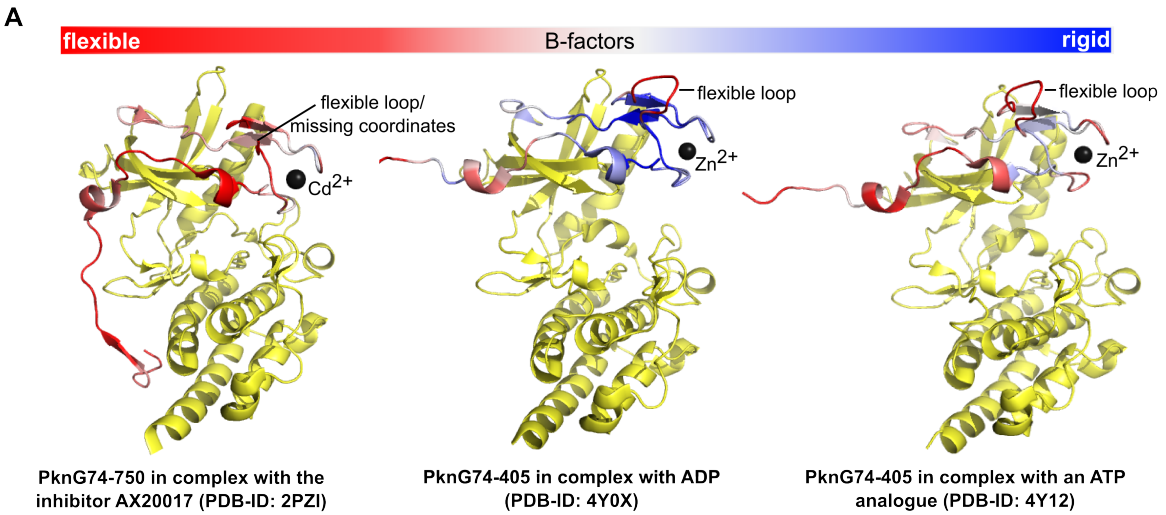
Wittwer et al.
SI Fig. S4

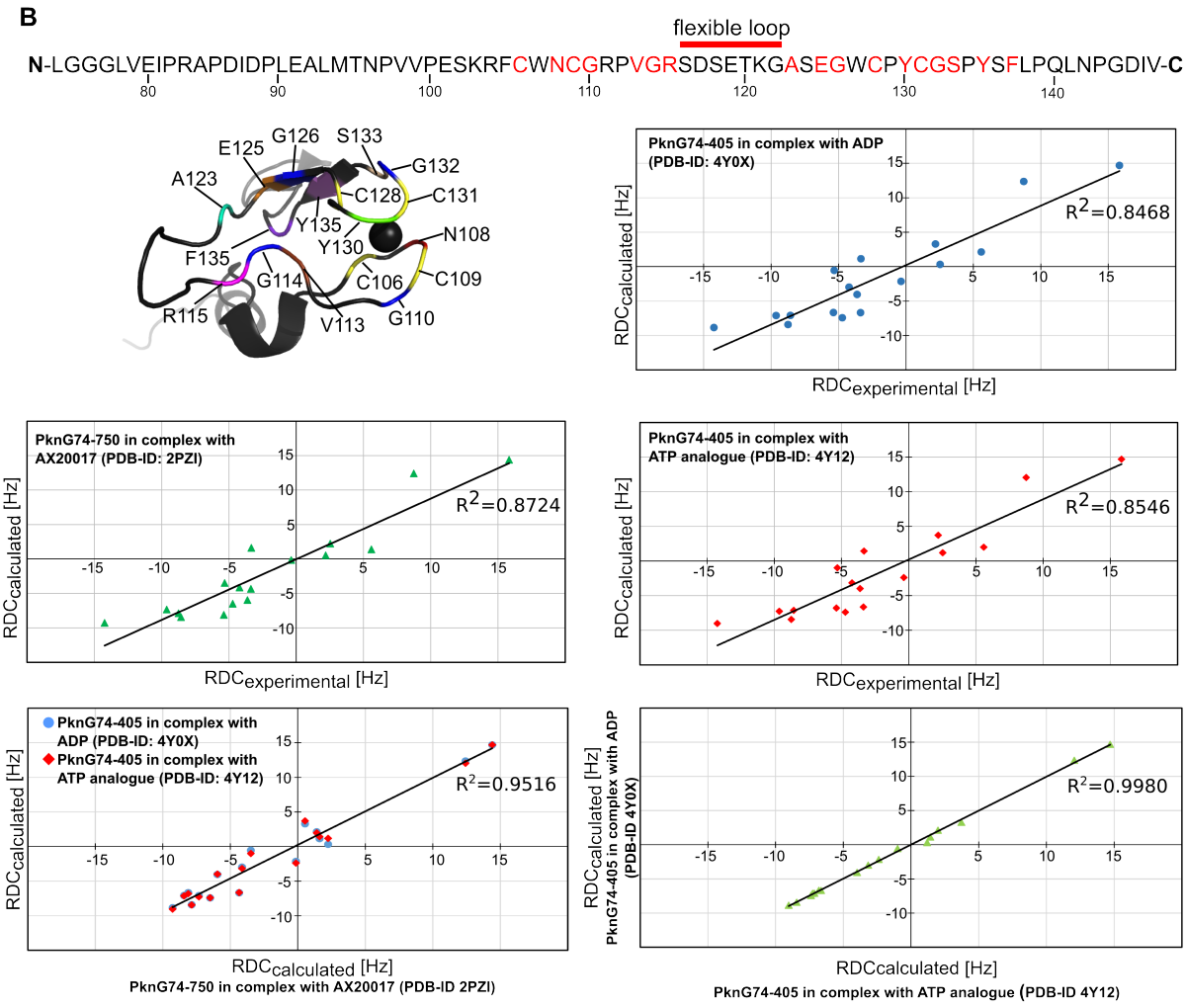


Wittwer et al.
SI Fig. S5

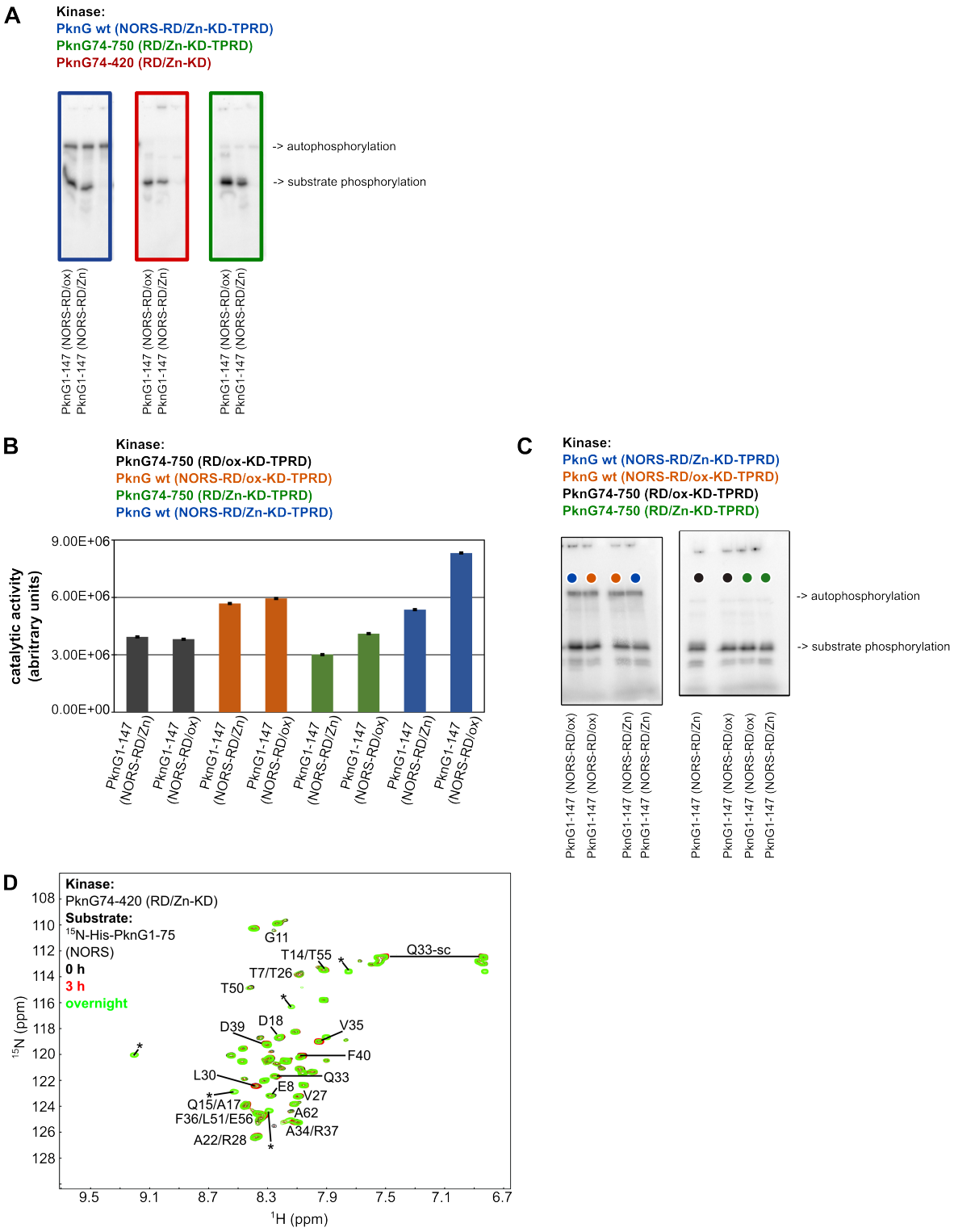


Wittwer et al.
SI Fig. S6

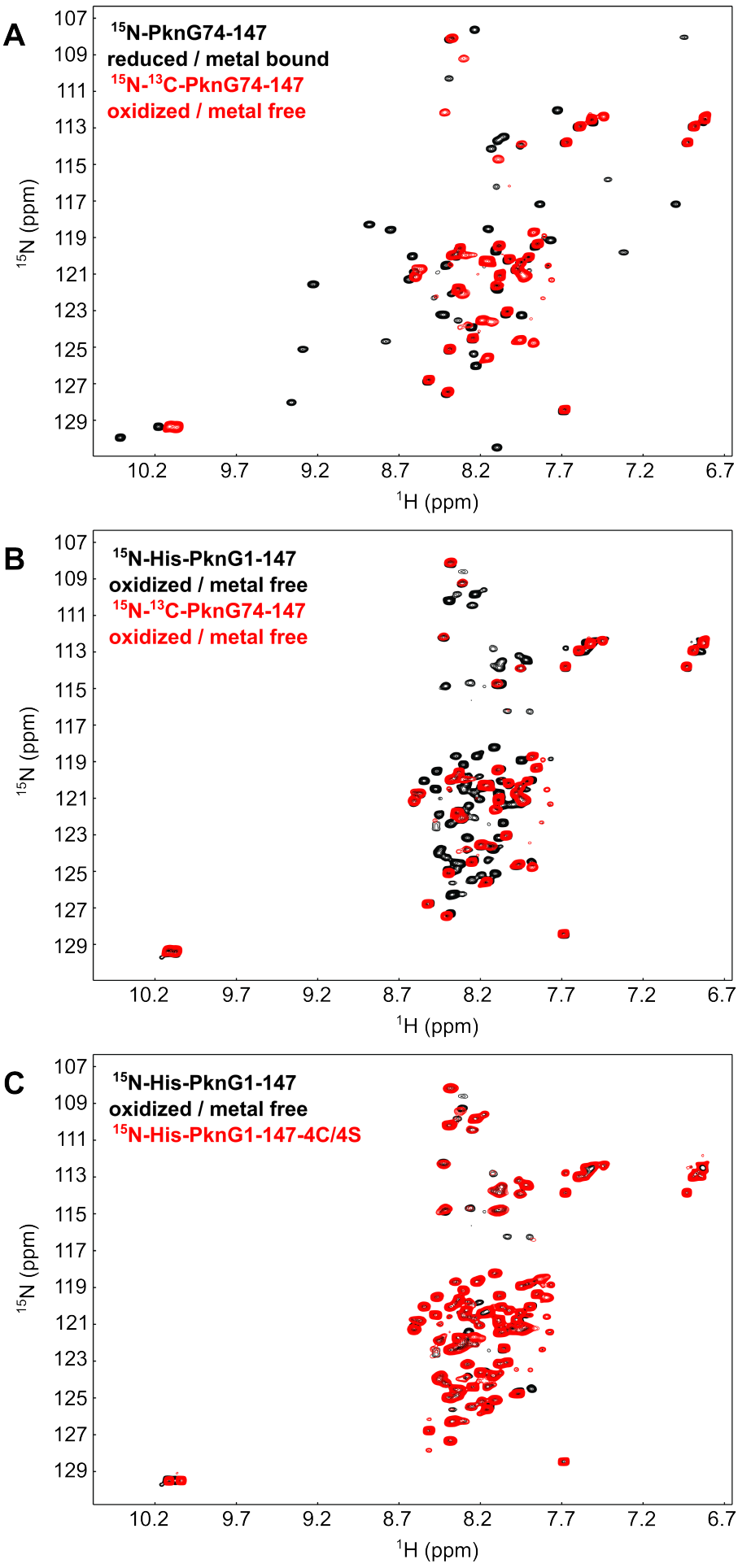




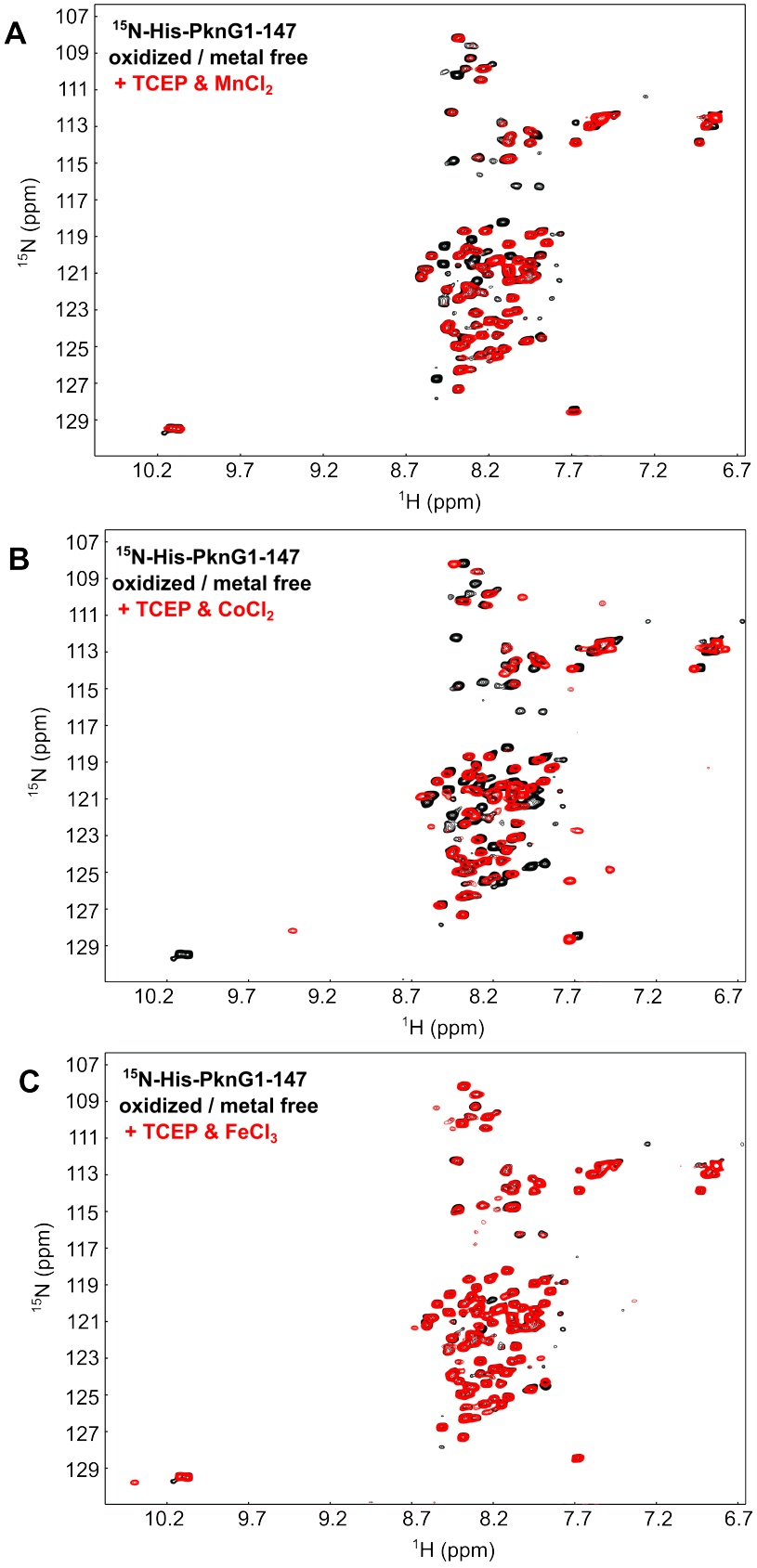
Wittwer et al.
SI Fig. S7



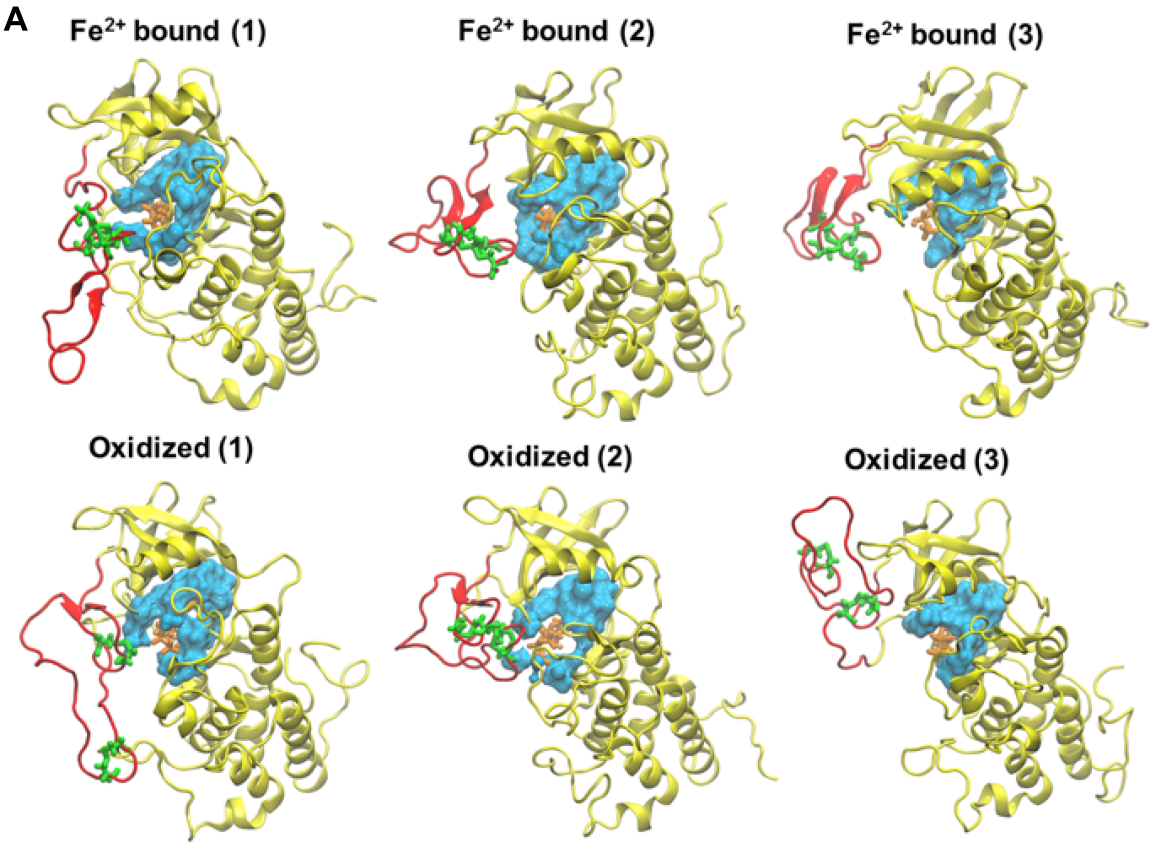
Wittwer et al.
SI Fig. S8

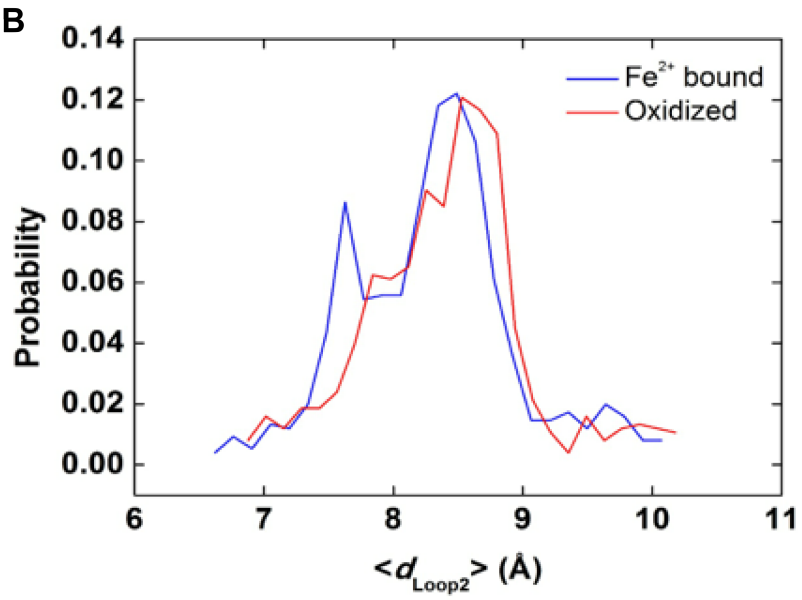


Wittwer et al.
SI Fig. S9

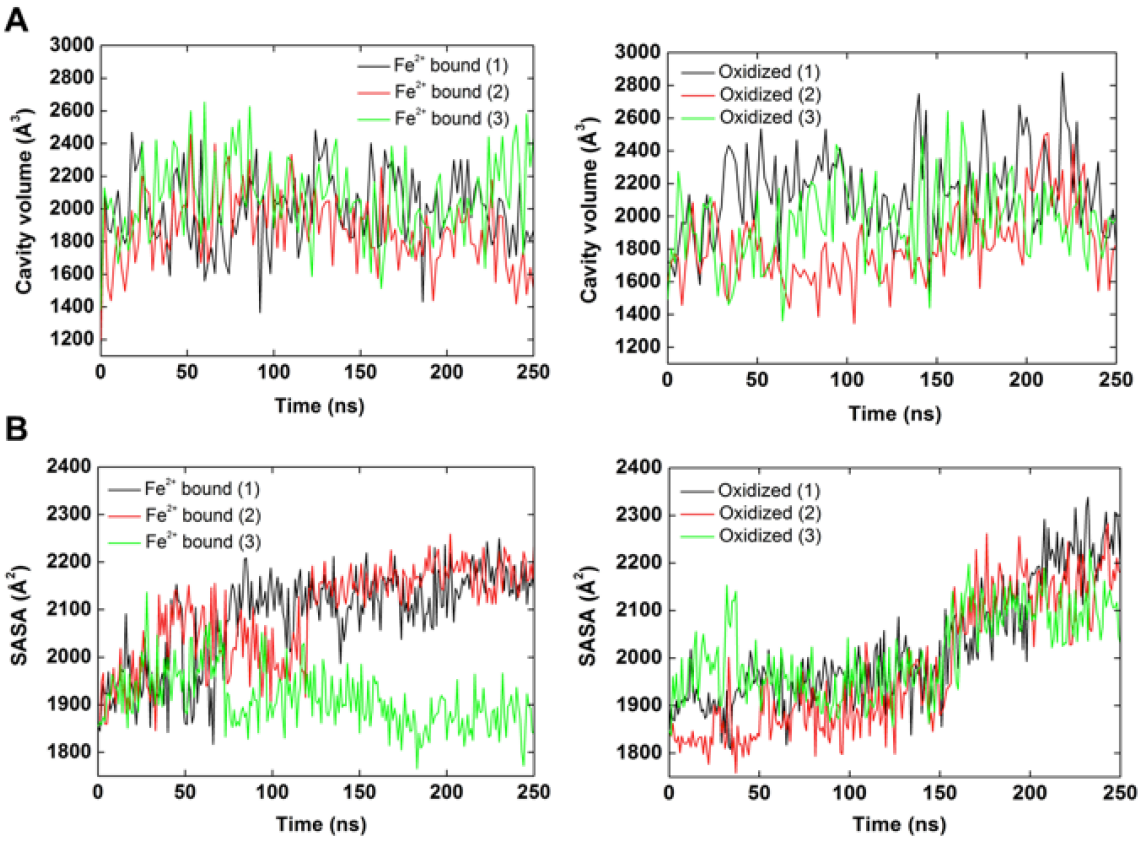


Wittwer et al.
SI Fig. S10





Wittwer et al.
SI Fig. S11



Wittwer et al.
SI Fig. S12