Erschienen in: FEBS Letters; 592 (2018), 7. - S. 1233-1245
https://doi.org/10.1002/1873-3468.13022

Structural characterization of the intrinsically disordered domain of *Mycobacterium tuberculosis* protein tyrosine kinase A

Anna Niesteruk[1], Marie Hutchison[1], Sridhar Sreeramulu[1], Hendrik R.A. Jonker[1], Christian Richter[1], Rupert Abele[2], Christoph Bock[2] and Harald Schwalbe[1]

1 Goethe-University Frankfurt am Main, Institute for Organic Chemistry and Chemical Biology, Centre for Biomolecular Magnetic Resonance (BMRZ), Frankfurt am Main, Germany
2 Goethe-University Frankfurt am Main, Institute of Biochemistry, Biocenter, Frankfurt am Main, Germany

Correspondence
H. Schwalbe, Goethe-University Frankfurt am Main, Institute for Organic Chemistry and Chemical Biology, Centre for Biomolecular Magnetic Resonance (BMRZ), Max-von-Laue-Straße 7, D-60438, Frankfurt am Main, Germany
Fax: +49 69 798 29515
Tel: +49 69 798 29737
E-mail: schwalbe@nmr.uni-frankfurt.de

(Received 16 February 2018, accepted 22 February 2018, available online 25 March 2018)

doi:10.1002/1873-3468.13022

Edited by Christian Griesinger

Although intrinsically disordered proteins or protein domains (IDPs or IDD) are less abundant in bacteria than in eukaryotes, their presence in pathogenic bacterial proteins is important for protein-protein interactions. The protein tyrosine kinase A (PtkA) from *Mycobacterium tuberculosis* possesses an 80-residue disordered region (IDD_{PtkA}) of unknown function, located N-terminally to the well-folded kinase core domain. Here, we characterize the conformation of IDD_{PtkA} under varying biophysical conditions and phosphorylation using NMR-spectroscopy. Our results confirm that the N-terminal domain of PtkA exists as an IDD at physiological pH. Furthermore, phosphorylation of IDD_{PtkA} increases the activity of PtkA. Our findings will complement future approaches in understanding molecular mechanisms of key proteins in pathogenic virulence.

Keywords: intrinsically disordered protein; *Mycobacterium tuberculosis*; nuclear magnetic resonance spectroscopy; protein denaturation; protein phosphorylation; protein tyrosine kinase

Intrinsically disordered proteins (IDPs) are characterized by a lack of stable structure and exist as an ensemble of heterogeneous conformers rapidly fluctuating in solution [1–4]. IDPs can adopt a variety of conformational states ranging from collapsed IDPs (molten globule, MG) to extended IDPs (native coil or native pre-molten globule, preMG) [5,6]. IDPs are highly abundant in various organisms and they play an important role during cellular regulation [7]. Approximately, 3% of proteins in bacteria are completely disordered. With regards to long disordered segments (> 41 residues) this value reaches 20%, which is comparatively lower to the 43% observed in eukaryotes [8,9]. In this report, we have investigated a pathogenic protein containing a 27% disordered N-terminal stretch. Since this protein has to function both in the bacterium and in the human host, the functional role of the disordered domain might be exerted within the host eukaryotes. Many IDPs are related to human diseases and are therefore particular interesting targets for drug-development [8,10,11]. Interactions of IDPs with ligands or binding proteins, post-translational modifications (PTMs, phosphorylation or glycosylation) or coordination of various counterions or osmolytes can promote conformational changes including disorder-to-order transitions

Abbreviations
CSP, chemical shift perturbation; eSTPK, eukaryotic-like serine/threonine protein kinase; IDD, intrinsically disordered domain; IDP, intrinsically disordered protein; KCD, kinase core domain; MG, molten globule; MptpA, low molecular weight protein tyrosine phosphatase A; PreMG, pre-molten globule; PtkA, protein tyrosine kinase A; PTM, post-translational modification; TEV, tobacco etch virus.

[5,12–14]. The high structural flexibility, the plasticity, and the dynamic behavior of IDPs play a central role in their biological function. These properties allow a combination of low affinity and high specificity during binding, potentially promoting interactions with more than one biological partner [12,15,16]. The evolutionary conserved features of IDPs are, however, a challenge for the probing of their structural, dynamical, and functional propensities; all the characteristics which are essential for establishing a structure-function-relationship and for rational drug design. NMR spectroscopy offers a wide range of methodologies for the extensive structural characterization of IDPs [15–19]. Despite the recent improvements of NMR techniques, allowing for better resolution and higher sensitivity, defining the conformational ensemble of IDPs sampled in the solution is still a challenge.

The *Mycobacterium tuberculosis (Mtb)* protein tyrosine kinase A (PtkA, 30.6 kDa) contains an N-terminal disordered region of 80 amino acids (intrinsically disordered domain, IDD) connected to the rigidly folded catalytic domain (kinase core domain, KCD). This structural arrangement is unique for bacterial kinases and is supposed to have a potential functional importance in the context of *Mtb* virulence. PtkA which is exclusively found in *Mtb* is a nonconventional tyrosine kinase with autophosphorylation activity [20]. PtkA interacts with the low molecular weight protein tyrosine phosphatase A (MptpA, 17.5 kDa), the critical virulence factor of *Mtb* [21–23]. The exact role of PtkA alone or in combination with its cognate phosphatase MptpA in promoting virulence or pathogenicity of *Mtb* is poorly understood. PtkA undergoes PTM (phosphorylation) in *Mtb* by presenting itself as a substrate for endogenous eukaryotic-like serine/threonine protein kinases (eSTPKs) [24]. In general, PTMs are known to play essential roles in the regulation of IDPs and may influence significantly the modulation of conformational transitions of IDDs.

We have previously determined the NMR solution structure of MptpA [25]. In this successive study, we characterize the conformational behavior of the IDD from the complementary kinase, PtkA (IDD_{PtkA}), under various conditions and describe the structural changes induced by phosphorylation. Our results indicate that IDD_{PtkA} exists as an unstructured state under native-like conditions. Moreover, we observe that the conformational behavior of the IDD in isolation is nearly identical to the full-length protein. Interestingly, denaturation studies of IDD_{PtkA} by pH and GdmCl delineate regions of nonrandom behavior in a consistent manner. Furthermore, we show that phosphorylation of the IDD has a regulatory effect on the enzymatic activity of PtkA.

Materials and methods

Protein expression and purification

The expression vector pET151/D-TOPO (Invitrogen, Carlsbad, CA, USA) including the His_6-tagged tobacco etch virus (TEV) PtkA sequence was provided from the laboratory of Y. Av-Gay (University of British Columbia). The DNA plasmid representing N-terminal PtkA domain the PtkA 1–81_opt_pET-166 was derived from the Gen Script manufacture. The plasmids were transformed into *Escherichia coli* BL21 (DE3) cells for the expression. The unlabeled protein was expressed using LB growth medium and the uniformly ^{15}N- or ^{13}C,^{15}N-labelled protein was expressed in M9 minimal medium using $^{15}NH_4Cl$ (1 g·L^{-1}) and ^{13}C-glucose (1 g·L^{-1}) as the sole nitrogen and carbon source, respectively. The growth media were supplemented with 1 mM ampicillin and cells were inoculated at 37 °C until an OD_{600} of 0.6 was reached. The protein expression was induced with 1 mM IPTG after 15 min of incubation on ice-water. The recombinant PtkA was expressed for over 16 h with aeration (120 r.p.m.) at 16 °C. The cells were harvested by centrifugation (4000 ***g***, 45 min, 4 °C). The cell pellet was either flash frozen and stored at −80 °C for later use or resuspended in lysis buffer (PBS-buffer, pH 7.5) supplemented with one EDTA-free protease inhibitor tablet (Roche, Mannheim, Germany). The cells were disrupted for 15 min using M-110P Microfluidizer (15 000 PSI) and the cell lysate was centrifuged (16 000 ***g***, 40 min, 4 °C) to separate the soluble fraction from the cell debris. The supernatant was loaded using loading buffer (pH 8.0) to a 5 mL Ni-NTA HisTrapHP column (GE Healthcare, Uppsala, Sweden). The purification was performed according to the manufacturer's guidelines. The His_6-tag was cleaved by addition of TEV protease overnight during the dialysis at 4 °C in dialysis buffer (pH 8.0) and separated on the Ni-NTA column using elution buffer (pH 8.0). Subsequently, preparative size exclusion chromatography was performed on a HiLoad 26/60 Superdex 75 column (GE Healthcare, Uppsala, Sweden) in NMR buffer (pH 7.5). The presence of protein was confirmed by SDS/PAGE analysis. The fractions containing pure protein were pooled, flash frozen and stored at −80 °C or immediately used for further experimental procedures.

Luciferase assay

The autophosphorylation activity of PtkA was determined using a Kinase-Glo® luminescent kinase assay (Promega, Madison, WI, USA; assay protocol see Materials and

Methods, Appendix S1). The luminescence was measured using VeritasTM Microplate Luminometer.

Phosphorylation of PtkA

Phosphorylation of the ^{15}N-labeled protein (full-length PtkA or KCD_{PtkA} or IDD_{PtkA} construct) was monitored at 298 K on a 600 MHz spectrometer. PtkA (300 μM) resolved in 180 μL of 50 mM HEPES/NaOH buffer (pH 7.5) containing 300 mM NaCl, 10 mM $MgCl_2$, 10 mM DTT, 10 mM ATP and 10% D_2O/90% H_2O was transferred to the 3 mm NMR-tube. After addition of PKA (4 μM) to the NMR-tube the PtkA phosphorylation was monitored using 2D-(^{1}H,^{15}N)-HSQC experiments recorded every 40 min within 1 day. 30 μL aliquot of the sample were taken for the MS analysis.

NMR spectroscopy

The NMR experiments were performed at 298 K on Bruker spectrometers (either 600-, 700-, 800-, 900-, or 950 MHz) equipped with TXI-HCN cryogenic probe. The spectrometers were locked on D_2O and 0.3 mM 2, 2-dimethyl-2-silapentane-5-sulfonic acid was used as an internal standard for spectral referencing. The acquisition and processing of the NMR data was carried out with TopSpin version 3.2 (Bruker Biospin) and analyzed using Sparky version 3.114. All experiments were performed on 300 μM protein samples, either ^{15}N or ^{15}N, ^{13}C labeled, in NMR buffer (pH 7.5) using 3 mm NMR-tube. The protocols for the backbone assignment, secondary chemical shift analysis, pH- and GdmCl- titration and the temperature series are provided in the Materials and Methods, Appendix S1.

Further protocols on the phosphorylation of unlabeled full-length PtkA, SEC-MALS, CD and buffer compositions are provided in the Materials and Methods, Appendix S1.

Results

Backbone assignment of IDD_{PtkA} (1–81)

The suitability of full-length PtkA for NMR structural characterization was investigated. The 2D-(^{1}H,^{15}N)-HSQC spectrum of PtkA (Fig. 1A) revealed considerable signal overlap characteristic for a random-coil region in addition to signals with a broad chemical shift dispersion indicative for a folded region. These data suggest that PtkA consists of unfolded regions together with a well-folded core domain. Typically, intrinsically disordered regions exhibit a strong bias in their amino acid sequence. By applying IUPred [26], a tool to predict unstructured regions in proteins, we found a good correlation between the NMR observed presence of unstructured regions and the prediction. IUPred predicts the first 80 amino acids in the N-terminal region of PtkA to be unstructured (Fig. S1B). The CD spectrum of IDD_{PtkA} indeed also shows a random-coil behavior (Fig. S2). These characteristics hamper the unambiguous backbone assignment of the N-terminal domain (Met^1-Leu^{81}) in PtkA (Fig. 1B). Also, the high abundance of proline residues (13 prolines) and the repetition of the same amino acid patterns complicate the resonance assignment of IDD. We decided to focus our investigations of the isolated IDD_{PtkA} (Met^1-Leu^{81}, 8.5 kDa). At a pH of 7.5 and 298 K, the 2D-(^{1}H,^{15}N)-HSQC spectrum of the IDD_{PtkA} shows 49 broad signals which corresponds to ~70% of the expected amide resonances (Fig. 1C). The signal line broadening and missing amide cross-peaks could arise either due to the pH, temperature, conformational exchange, solvent exchange or protein oligomerization. In order to assess the pH-effect we first performed a pH titration. We varied the pH value from 7.5 to 2.0, which resulted in a narrowing of the amide signals. The 2D-(^{1}H,^{15}N)-HSQC spectrum of IDD_{PtkA} at pH 2.0 shows 69 well-resolved cross-peaks (Fig. 1C) in the chemical shift region typical for a random-coil protein.

Under these conditions, we were able to assign all the backbone atoms using a set of 3D-NMR experiments (HNCO, HNCACB, (H)N(CA)NNH, HNHA and ^{1}H,^{15}N-NOESY). The backbone assignments from pH 2.0 were transferred to pH 7.5 by using the pH series of 2D-(^{1}H,^{15}N)-HSQC experiments. Forty seven backbone amide signals of IDD_{PtkA} visible at native-like conditions (pH 7.5) were assigned. The above observations indicate that the IDD_{PtkA} is unstructured under these conditions. To rule out the possibility of oligomerization of IDD_{PtkA}, we performed a series of 2D-(^{1}H,^{15}N)-HSQC experiment at concentrations ranging between 300 and 35 μM. We used the proton line-width and the peak intensity at each concentration as reporters (Fig. S3). Our analysis shows no concentration-dependent reduction in the line-width. Moreover, the normalized cross-peak intensity in the series of 2D-(^{1}H,^{15}N)-HSQC spectra are essentially the same, indicating no oligomerization of IDD_{PtkA}.

Secondary structure composition of IDD_{PtkA}

2D-(^{1}H,^{15}N)-HSQC based backbone analysis suggests similar conformational properties of the IDD in isolation and in full-length PtkA. An overlay of the 2D-(^{1}H,^{15}N)-TROSY/HSQC spectra of full-length PtkA and IDD_{PtkA} measured at pH 7.5 and 298 K shows that the chemical shifts are essentially the same, except for the residues located proximate to the linker region between the IDD and the KCD of PtkA as well as for

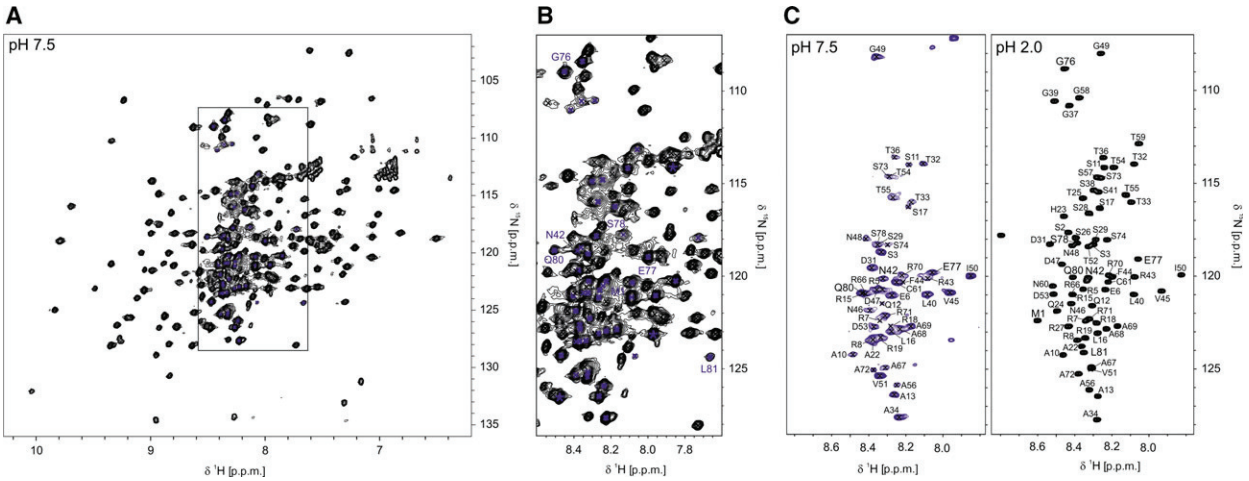

Fig. 1. Backbone assignment of IDD_{PtkA}. (A) 2D-(^{1}H,^{15}N)-TROSY spectrum of the ^{2}H, ^{15}N, ^{13}C labeled full-length PtkA at pH 7.5 and 298 K. Spectral region corresponding to the IDD is framed and the resonances assigned to the IDD are labeled in blue. (B) the enlargement of the overcrowded IDD-region between 7.6 and 8.6 p.p.m. in the ^{1}H- and 107 and 129 p.p.m. in the ^{15}N-dimension of the 2D-(^{1}H,^{15}N)-TROSY spectrum of the full-length PtkA. (C) 2D-(^{1}H,^{15}N)-HSQC spectra of the ^{15}N-labeled IDD_{PtkA} at pH 7.5, 298 K (blue, left) with tentative assignment transferred based on pH-titration series from assignment generated at pH 2.0, 298 K (black, right). The backbone assignment of PtkA has been deposited in the Biological Magnetic Resonance Data Bank (BMRB) with entry number 34204.

Asn^{48} which could suggest a week interaction between the domains (Fig. S4). Since the backbone amides are prone to chemical exchange with varying pH, we additionally performed a residue specific analysis of the nonexchangeable carbon chemical shifts (Cα, Cβ and CO) for the isolated IDD_{PtkA} at pH 2.0 and the full-length PtkA at pH 7.5.

The secondary chemical shift (Δδ) of a certain protein nucleus was calculated as the difference between the observed chemical shift and the corresponding 'random-coil' value (Fig. 2). The relationship between the Δδ and the secondary structure elements allows us to distinguish between structured (α- helices or β-strand) and unstructured (random-coil) protein regions [27]. Our results show as expected, a distinct correlation for the large Δδ values of the KCD, indicative for the existence of the secondary structural elements for this core domain. On the other hand, the very small Δδ values for the IDD evidence that there are no clear secondary structural elements, suggesting a random-coil state for this N-terminal domain. In addition, we performed a combined chemical shift analysis using Delta2D [28] and SSP [29]. Both methods clearly illustrate (Fig. S6) the well-defined KCD and the disordered state of the IDD.

Structural behavior of IDD_{PtkA} under different biophysical conditions

Intrinsically disordered proteins lack defined secondary structural elements under native-like conditions but can adapt a stable fold in a stimulus-dependent manner. Protein-protein interaction, change in pH, temperature or PTMs (phosphorylation) can induce structural changes in the IDPs thereby altering the function of the protein or help during translocation [12,30–32]. We therefore probed the effect of pH, chemical denaturant (guanidine hydrochloride, GdmCl) and phosphorylation on the IDD_{PtkA}.

Effect of pH on IDD_{PtkA}

We monitored the effect of pH (ranging from 7.5 to 2) on the IDD_{PtkA} using 2D-(^{1}H,^{15}N)-HSQC spectra. At native-like condition (pH 7.5), 68% of the amide cross-peaks of the IDD_{PtkA} are visible. Missing signals correspond to the amino acid region of Met^{1}-Ser^{3}, Arg^{7}, Ser^{17}, His^{23}-Ser^{29}, Gly^{37}-Gly^{39}, Ser^{57}-Asn^{60}, Gly^{76} and Leu^{81}. When lowering the pH from 7.5 to 5.5, 20 (30%) additional signals can be observed. A further decrease in the pH to 2.0 resolves all of the 67 (100%) expected amide cross-peaks (Fig. 3). Beside the (dis)appearance of amide cross-peaks, chemical shift perturbations (CSPs) can be observed due to the change in pH. The strongest CSPs were observed for the residue Asp^{31} and the region Asn^{42}-Ala^{56} of the IDD_{PtkA} (Fig. 3B). These residues are adjacent to the regions (His^{23}-Ser^{29}, Gly^{37}-Gly^{39}, Ser^{57}-Asn^{60}) of which the amide cross-peaks are missing at pH 7.5. Chemical shift analysis using the best-fit method shows, that most of the resonances move along a line with changing pH. A slight deviation is observed for

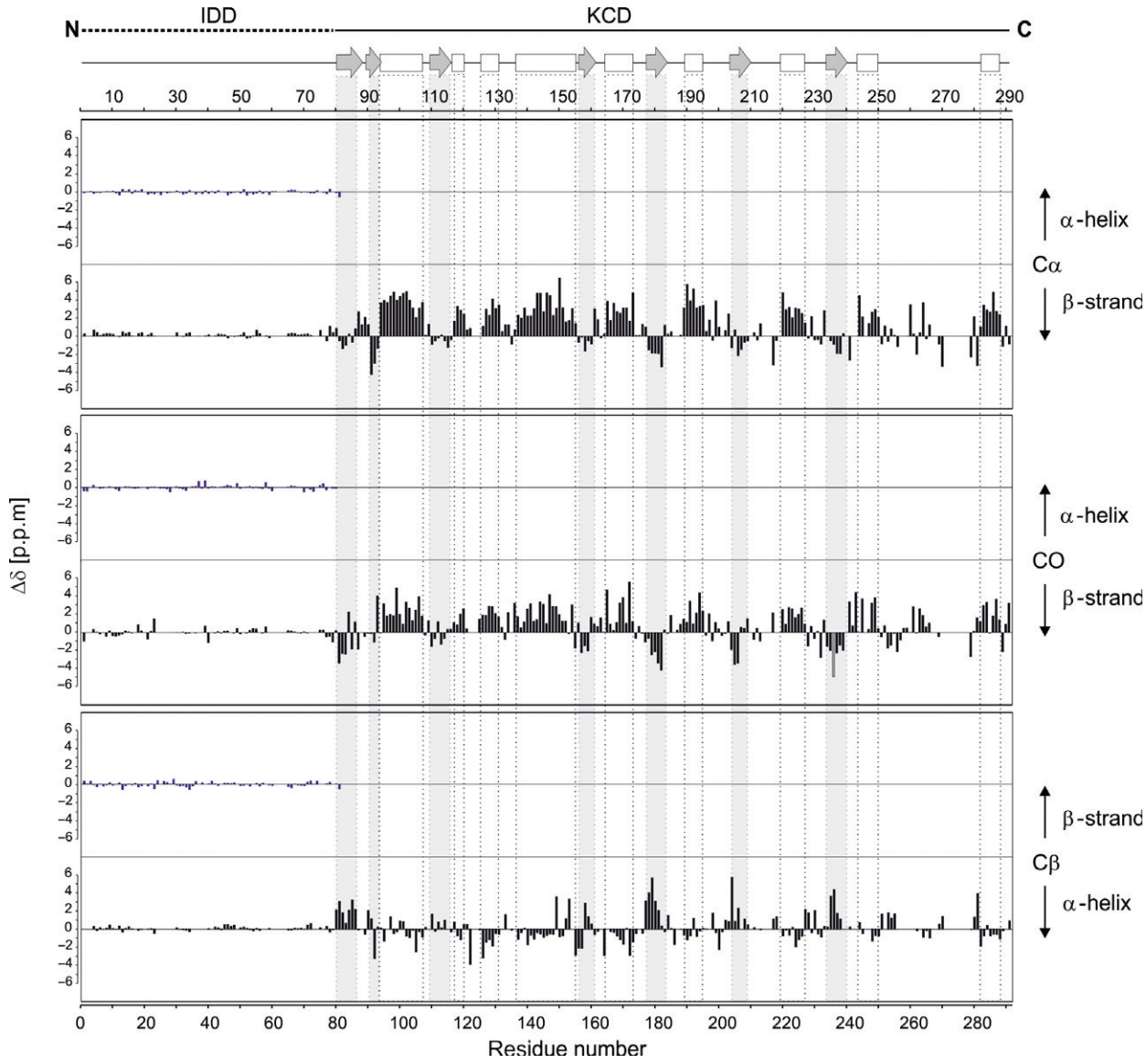

Fig. 2. Comparison of secondary chemical shifts of PtkA. Secondary chemical shifts (Δδ, p.p.m.) of IDD_{PtkA} at pH 2.0 (blue) and full-length PtkA at pH 7.5 (black): ΔδCα (top), ΔδCO (middle) and ΔδCβ (bottom). Δδ = δexp − δrc, with δexp = experimental chemical shift and δrc = random coil chemical shift generated using a JAVA-script from Alex Maltsev, available on the website of the University of Copenhagen (http://www1.bio.ku.dk/english/research/bms/research/sbinlab/groups/mak/randomcoil). Positive Δ(δCα) and Δ(δCO) values indicate α-helix, while negative β-strand. The trend of Δ(δCβ) is opposite: negative values indicates β-strand and positive α-helix. For the expanded views of the secondary chemical shifts of the IDD see Fig. S5.

Asp^{31}, Asp^{47}, and Asp^{53} in the ^{15}N-dimension and Ser^{11}, His^{23}, Asp^{31}, Asp^{47}, Asp^{53}, and Ala^{56} (falling under the line of best-fit) and Thr^{55}, Thr^{33}, Thr^{54}, and Gly^{49} (lying above the best-fit line) in the ^{1}H-dimension (Fig. S7). This could be a consequence of different pH-induced effects per residue type or fast exchange between two states, possibly via an intermediate state. In summary, we observe chemical shift changes of the signals together with line-broadening and disappearing peaks during the titration, showing the pH-dependent dynamics of IDD_{PtkA}.

Effect of GdmCl on IDD_{PtkA}

In addition to the pH studies, we also investigated the influence on the IDD_{PtkA} due to chemical induced denaturation. A set of titration experiments combining one or more non-native conditions were

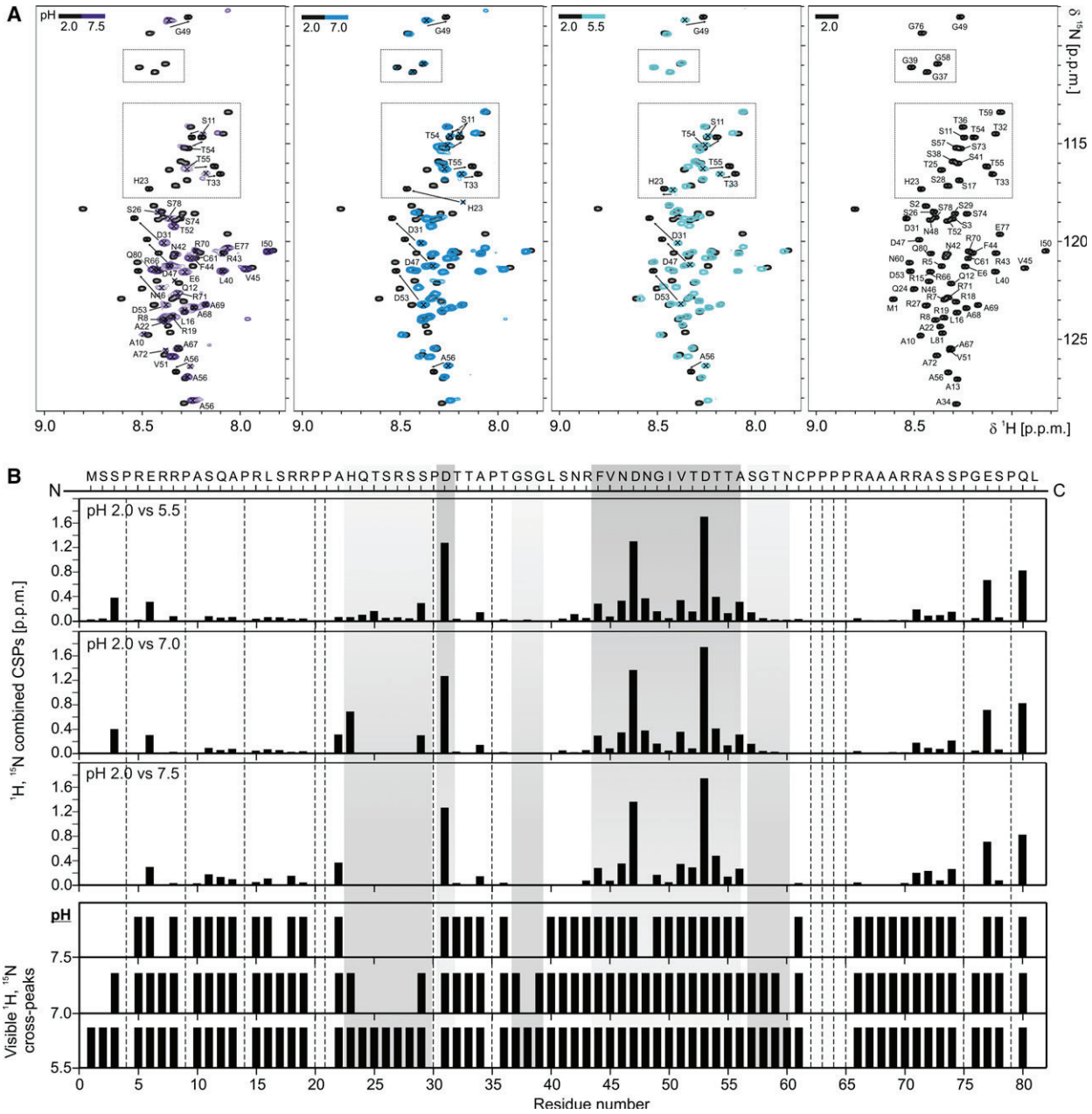

Fig 3. pH effect on the backbone amide groups of IDD_{PtkA}. (A) Overlay of the 2D-(^{1}H,^{15}N)-HSQC spectra of IDD_{PtkA} at pH 7.5, pH 7.0, pH 5.5 vs pH 2.0. Signals which are mostly affected by the pH changes are labeled with the corresponding amino acid. The shift direction is indicated by the arrows. The regions most affected by the pH-induced acidification are highlighted with boxes. (B) ^{1}H, ^{15}N combined CSPs obtained after analysis of 2D-(^{1}H,^{15}N)-HSQC spectra of IDD_{PtkA} at pH 2.0 vs pH 5.5 (top), pH 2.0 vs pH 7.0 (middle) and pH 2.0 vs pH 7.5 (bottom). Dotted-lines corespond to the proline residues; bars highlight residues, whose signals appears with the pH decrease (light grey) and the regions with significant CSPs (dark grey). Bars chart (bottom) shows schematically the amide cross-peaks of IDD_{PtkA} detected at each pH-value (pH 5.5, pH 7.0 and pH 7.5).

performed using the denaturating agent GdmCl. Denaturation of IDD_{PtkA} using increasing concentrations of GdmCl was monitored by 2D-(^{1}H,^{15}N)-HSQCs at pH 7.5 at two different temperatures: 298 K and 283 K (Fig. 4).

The spectra reveal that 80% of the signals observed at pH 7.5 (298 K) disappear in presence of 4 M GdmCl and reappear again almost completely at higher GdmCl concentration (5 M) as sharp signals, with a slightly different chemical shift (Figs 4 and 5). The most affected amide

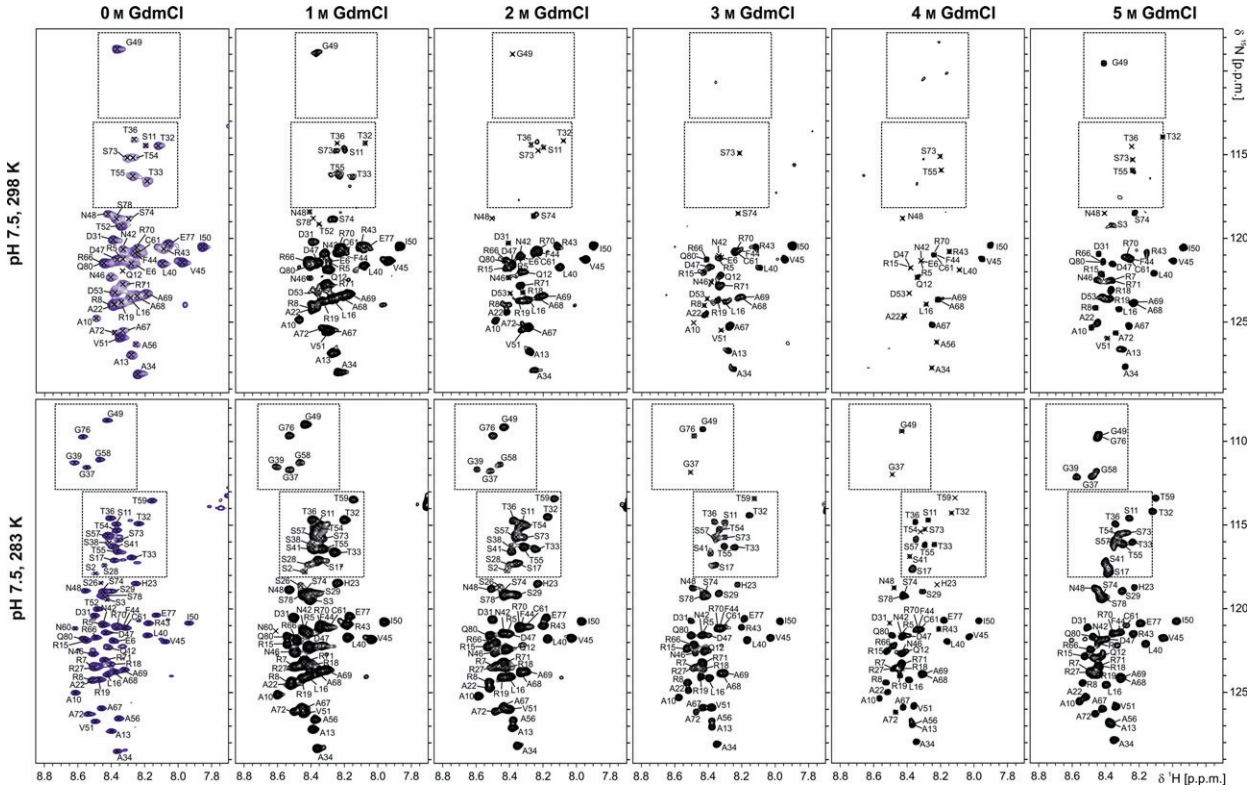

Fig 4. Treatment of IDD_{PtkA} with GdmCl at pH 7.5. 2D-(^{1}H,^{15}N)-HSQC spectra of IDD_{PtkA} at pH 7.5 and 298 K (top) and 283 K (bottom), recorded during the titration with GdmCl from 1 M up to 5 M, with 1 M increments. The regions which are most affected by GdmCl are highlighted with boxes.

cross-peaks correspond to the residues glycine, serine and threonine. For a significant number of residues, the signal intensities increase at lower temperature (283 K) and their linewidths become narrower due to reduced exchange with solvent (Fig. 4). Furthermore, the IDD_{PtkA} denaturates completely upon titration with GdmCl at non-native pH 2.0, which further reduces the dispersion of the amide signals (Fig. 6). In contrast to the GdmCl-induced denaturation of IDD_{PtkA} at pH 7.5, no signal attenuation was observed. The chemical shift analysis using the best-fit method shows that most of the chemical shifts move along the line during the GdmCl-titration. However, slight deviation in the ^{1}H-dimension can be observed for Arg^{18}, Thr^{25}, Ser^{29}, Thr^{32}, Asn^{42}, Arg^{43}, Thr^{54} ,and Arg^{67} (Fig. 6B). Altogether, these results suggest that in the presence of denaturant the IDD_{PtkA} traverses from a native state to denatured state.

Effect of the temperature modulation on the backbone amide groups of IDD_{PtkA}

The temperature-dependence of the amide chemical shift can be used to identify those groups involved in hydrogen-bonding. Slowly exchanging amide protons, involved in hydrogen bonding, are characterized by a temperature coefficient more positive than -4.5 ppb·K^{-1} [33]. To analyze whether some residues of IDD_{PtkA} are involved in the hydrogen-bonding, we determined the amide temperature coefficients from a series of 2D-(^{1}H,^{15}N)-HSQC spectra recorded at 301 to 280 K (with 3 K increments). The obtained temperature coefficients range between -3.6 to -9.5 ppb·K^{-1} (Fig. S9). For most residues the temperature coefficients are more negative than -4.5 ppb·K^{-1}, indicating fast exchange with the solvent, except for the C-terminal part of the IDD_{PtkA} where the temperature coefficients are fluctuating. Noteworthy, significant signal attenuation was observed for the glycine residues with increasing temperature. At lower temperatures (280–283 K) all five glycine signals of the IDD_{PtkA} (Gly^{37}, Gly^{39}, Gly^{49}, Gly^{58}, and Gly^{76}) were visible in the 2D-(^{1}H,^{15}N)-HSQC spectrum, whereas at higher temperatures (298–301 K) only glycine (Gly^{49}) could be detected. Similar effects are also observed for the serine and threonine residues.

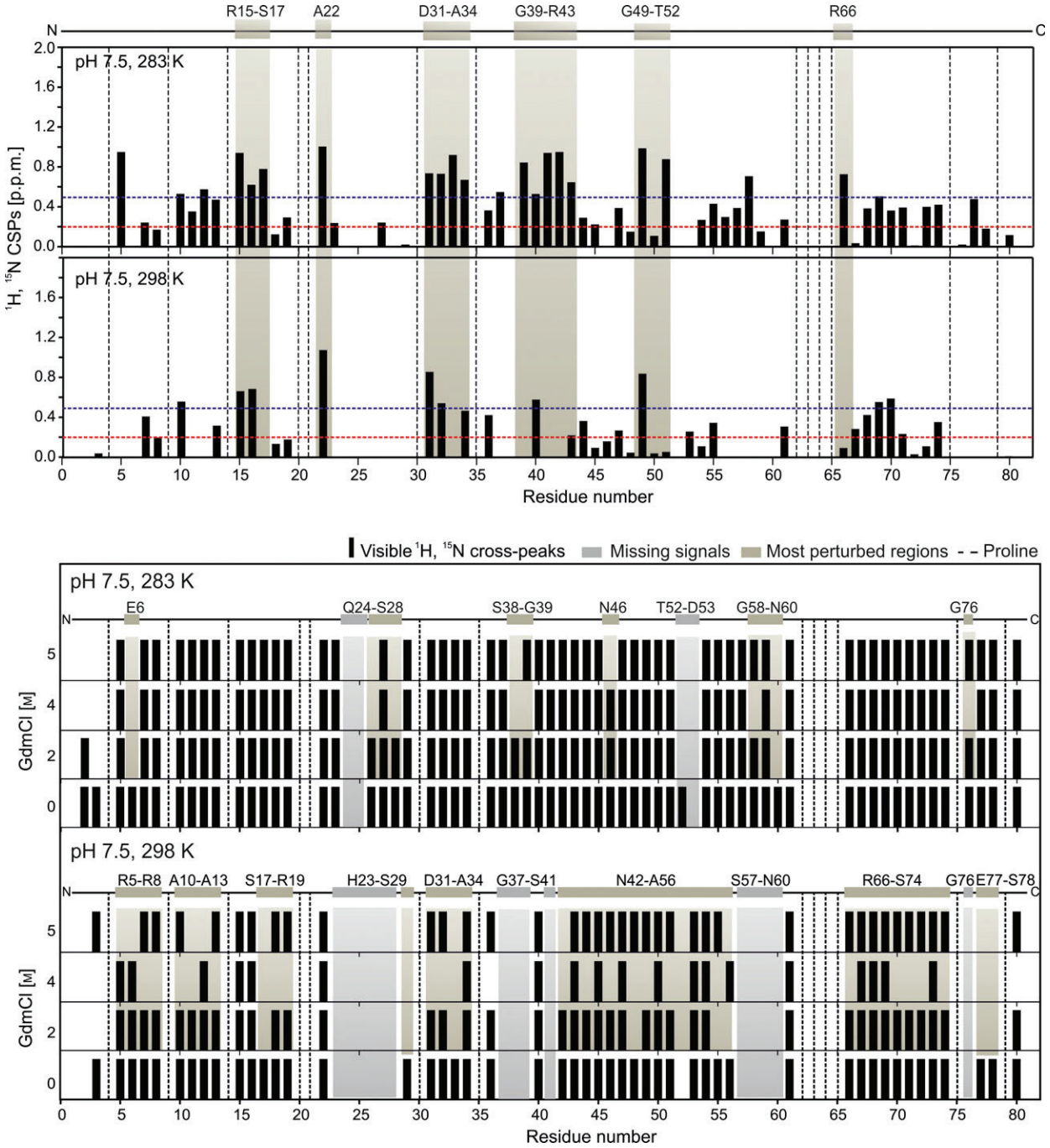

Fig 5. GdmCl effect on the backbone amide groups of IDD_{PtkA} at pH 7.5. (A) ^{1}H, ^{15}N combined CSPs obtained from analysis of 2D-(^{1}H, $z^{15}N$)-HSQC spectra of IDD_{PtkA} with 0 M vs 5 M GdmCl at pH 7.5, 283 K (top) and 298 K (bottom). Dotted-lines corresponds to the proline residues; regions with significant CSPs are highlighted in beige. (B) Bars chart showing schematically visible amide cross-peaks of IDD_{PtkA} pH 7.5 and 283 K (top) and 298 K (bottom), detected im presence of 0, 2, 4 and 5 M GdmCl.

Phosphorylation-induced conformational changes

Threonine-specific phosphorylation of PtkA by *Mtb* endogenus eSTPK was recently reported [34]. Previous studies by Husson *et al*. propose serine/threonine specific phosphorylation as an important regulatory mechanism of *Mtb* virulence [35]. We investigated the effects of the phosphorylation on the structural properties of the PtkA, KCD_{PtkA} and IDD_{PtkA} by NMR.

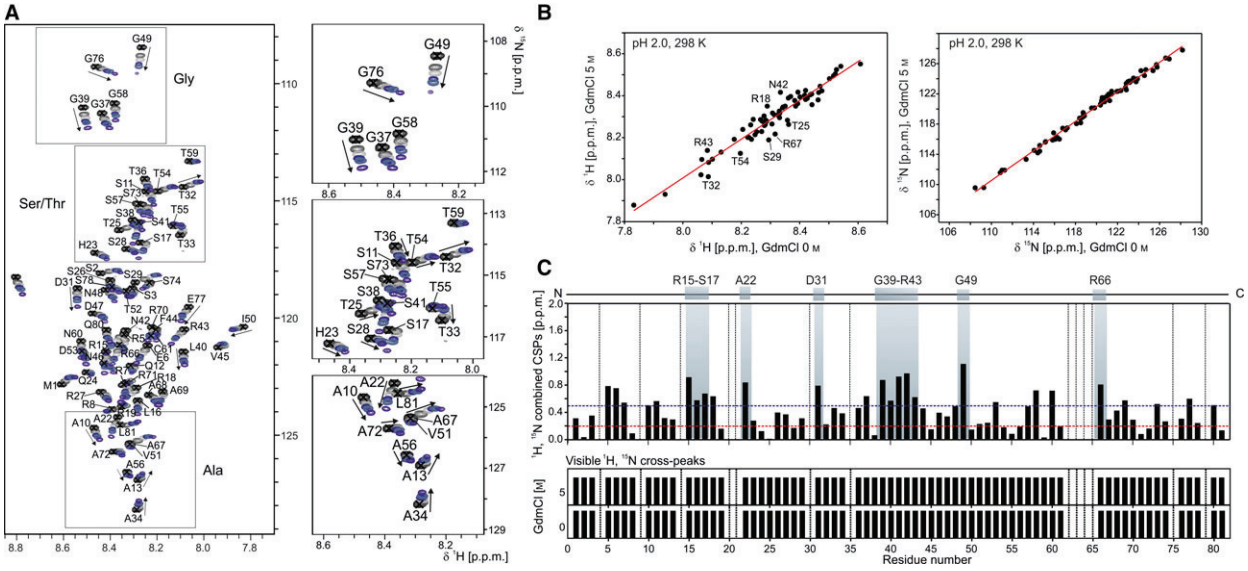

Fig 6. Combined pH and GdmCl effect on to the backbone amide groups of IDD_{PtkA}. (A) overlay of the 2D-(^{1}H,^{15}N)-HSQC spectra of IDD_{PtkA} at pH 2.0, 298 K, recorded upon addition of increasing amounts of GdmCl from 1 M up to 5 M, with 1 M increments (spectrum without GdmCl colored in black, spectrum with 5 M GdmCl collored in blue). On the right a zoom on the spectral regions of Gly, Ser/Thr and Ala is shown. (B) best-fit analysis showing the distribution of the chemical shift of IDD_{PtkA} at pH 2.0, 298 K with 0 M vs 5 M GdmCl in the ^{1}H-dimension (left) and ^{15}N-dimension (right). Signals showing deviations from the best-fit analysis are labelled to the corresponding residues. (C) ^{1}H, ^{15}N combined CSPs obtained after analysis of 2D-(^{1}H,^{15}N)-HSQC spectra of IDD_{PtkA} at pH 2.0, 298 K (top) with 0 M vs 5 M GdmCl. Dotted-lines correspond to the proline residues; blue bars highlight the regions with significant CSPs. Bar chart schematically showing visible amide cross-peaks of IDD_{PtkA} at pH 2.0 and 298 K without GdmCl (bottom) and with 5 M GdmCl (top).

The phosphorylation was performed using a eukaryotic serine/threonine kinase, protein kinase A (PKA). The 2D-(^{1}H,^{15}N)-HSQC spectrum of the full-length PtkA after incubation with ATP and PKA showed time-dependent phosphorylation of PtkA, reflected by the appearance of additional signals (Fig. S10A). MS analysis of full-length PtkA incubated for 24 h with ATP and PKA revealed three-fold phosphorylation of PtkA (Fig. S11). We conducted a similar phosphorylation study on the KCD_{PtkA} and IDD_{PtkA}. The primary amino acid sequence of the IDD is highly abudant in serine and threonine residues (13 serine/8 threonine) and therefore presents potential sites for PTM. Our data shows that PKA specifically phosphorylates the IDD_{PtkA} and not the KCD_{PtkA}. Furthermore, the amide chemical shifts of the newly appearing signals in the 2D-(^{1}H,^{15}N)-HSQC of the IDD_{PtkA} upon phosphorylation overlap with those observed in the full-length PtkA, indicating that the same residues were phosphorylated (Fig. S10B). Typically, phosphorylation of serine or threonine residues in IDPs results in a large downfield chemical shift changes of their backbone amide signals ($\Delta\delta \sim 0.5$/1.5 p.p.m. (^{1}H/^{15}N) at pH ~ 7.0) [36]. Based on the assignment of the nonphosphorylated protein and by following the disappearing signals upon phosphorylation, the substrate sites can be readily identified. The 2D-(^{1}H,^{15}N)-HSQC spectra of IDD_{PtkA} incubated for various time points with PKA and ATP were compared to the 2D-(^{1}H,^{15}N)-HSQC spectrum of the nonphosphorylated protein (Fig. 7). The amide cross-peaks of seven (Ser^3, Ser^{11}, Ser^{17}, Ser^{29}, Ser^{73}, Ser^{74}, Ser^{78}) of 13 serines, conserved in the sequence of IDD_{PtkA}, are visible and assigned at pH 7.5. Based on the signal dissappearance and appearance of new cross-peaks, two (Ser^{11} and Ser^{74}) of four phosphorylated serines could be identified. For the other two serines, we were not able to follow the same trend due to which they remain unassigned. Furthermore, spectra acquired after longer incubation time (5 days) generated an estimated 11 new additional signals. An overlay of the 2D-(^{1}H,^{15}N)-HSQC spectra of nonphosphorylated and phosphorylated IDD_{PtkA}, clearly shows that the upcoming of phosphorylated serines appear at typical random-coil positions [37] (Fig. 7B). Noteworthy, the amide signals of four of five glycines appear upon phosphorylation, which are otherwise only detectable at lower temperature. These glycines are in close proximity to serine residues, suggesting that phosphorylation of the IDD causes local changes

in the chemical environment and likewise the partial dynamics of this domain.

Phosphorylation of the IDD and PtkA activity

We monitored the catalytic activity of both phosphorylated and nonphosphorylated forms of PtkA using a luminescent kinase assay. Phosphorylated PtkA was four times more active than its nonphosphorylated form (Fig. 8). These results indicate that the phosphorylation of the IDD alters the structural architecture of PtkA, resulting in enhanced substrate binding and activity.

Discussion

In this study, we have characterized the N-terminal IDD_{PtkA} of the unique bacterial tyrosine kinase PtkA from *Mtb*, using NMR spectroscopy. Our results describe the conformational behavior of IDD_{PtkA} under native-like and various biophysical conditions. Chemical shift analysis reveals no persistent secondary structural elements, indicating that the IDD_{PtkA} is unfolded. Furthermore, the IDD undergoes slow structural fluctuations over different conformational states, either in full-length PtkA or when in isolation. This leads to extensive line-broadening of NMR signals. The amide

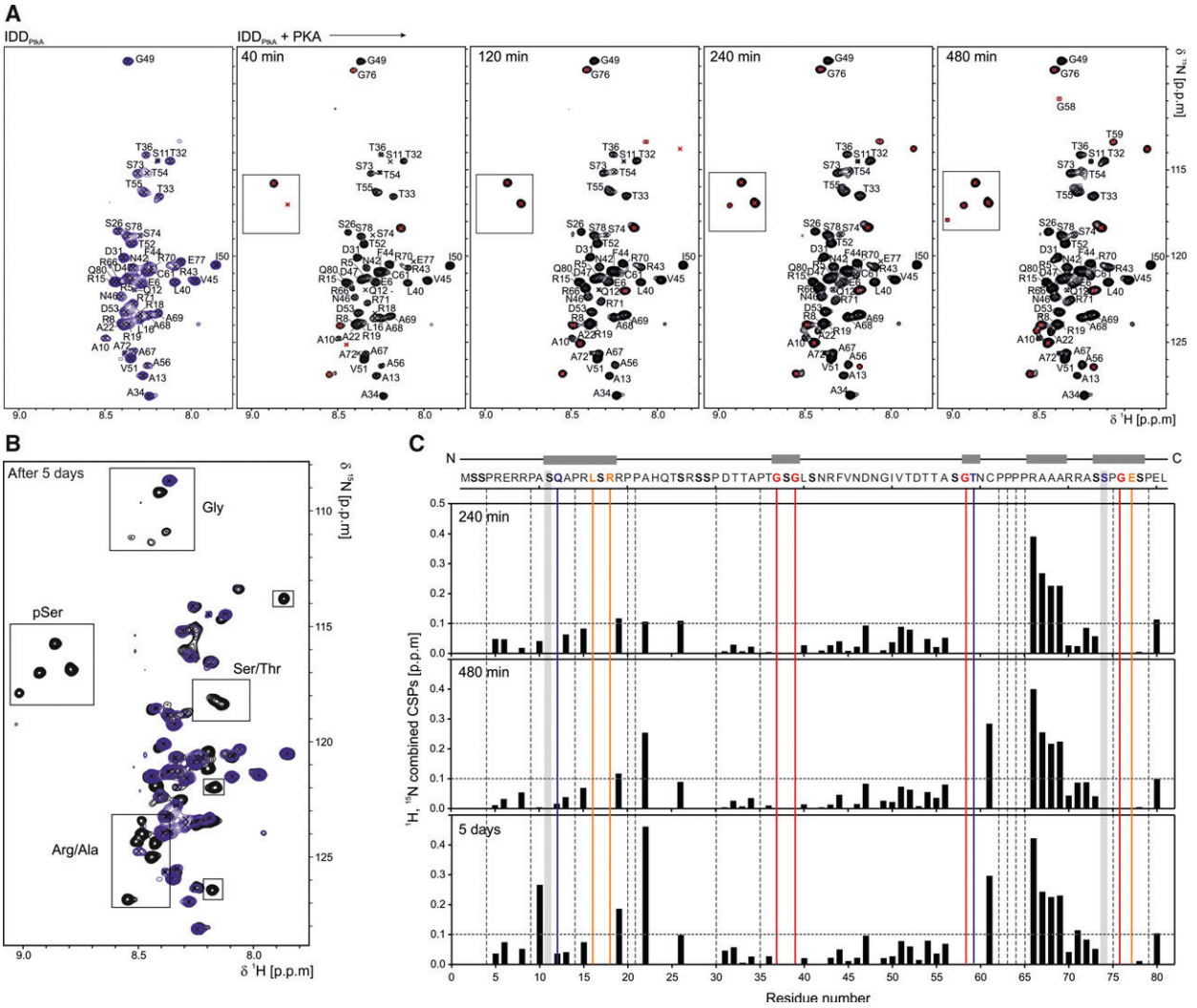

Fig. 7. Phosphorylation of IDD_{PtkA} by PKA. (A) 2D-(^{1}H,^{15}N)-HSQC spectra of IDD_{PtkA} without PKA (blue) and time resolved spectra after addition of PKA (black). New signals are labeled in red, signals indicating phosphorylation on serine are framed. (B) overlay of the 2D-(^{1}H,^{15}N)-HSQC spectra of IDD_{PtkA} without PKA (blue) and after 5 days incubation with PKA. Regions with new signals are framed. (C) ^{1}H, ^{15}N combined CSPs obtained after analysis of 2D-(^{1}H,^{15}N)-HSQC spectra of IDD_{PtkA} with and without PKA. Appearing signals (red), disappearing signals (orange), signals which disappear after addition of PKA and appear during the incubation (blue), prolines are indicated with dotted lines. The identified phosphorylated Ser^{11} and Ser^{74} are highlighted in grey.

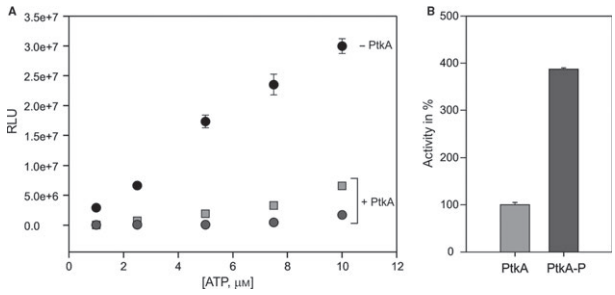

Fig. 8. Autophosphorylation activity of PtkA. (A) the autophosphorylation activity of PtkA measured at different ATP concentrations (1, 2.5, 5, 7.5 and 10 μM) using nonphosphorylated form of PtkA (light gray) and phosphorylated form of PtkA (dark gray). Decrease in the amount of emitted light (relative light units, RLU) indicate autophosphorylation ability of PtkA. (B) PtkA activity in % calculated based on the RLU values obtained from the Luciferase assay performed in presence of 10 μM of ATP. The activity of nonphosphorylated PtkA was set to 100%. The comparison of the nonphosphorylated and phosphorylated form of PtkA shows a 4-fold activity increase after phosphorylation.

solvent exchange properties differ per residue type and strongly correlate with its neighboring residue side chain [38]. Generally, in folded or partially folded proteins, the amide groups located on the surface exchange faster with the solvent than those buried and/or participating in hydrogen bonding. Modulation of the pH changes the charge distribution of the amino acid, thereby affecting the electrostatic interactions, which are essential for the integrity of the structure, dynamics and the function of the protein. Chemical- (GdmCl) and pH-induced denaturation of IDD_{PtkA} shows similar nonrandom coil like behavior at least for two regions of the IDD_{PtkA} (Fig. S12). Our results indicate that the pH-induced conformational changes within the IDD_{PtkA} are reversible and that the GdmCl-induced denaturation of IDD_{PtkA} progresses via an intermediate state. Even though the amino acid sequence determines the native conformation, the environment dictates the pathway it takes to fold and attain the physiologically relevant structural conformation.

In summary, our results consistently show that amino acid regions of IDD_{PtkA} corresponding to Arg^{15}-Ser^{17}, Asp^{31}, Gly^{37}-Arg^{43}, Asn^{48}-Thr^{52}, and Gly^{58}-Thr^{59}, undergo significant changes due to environmental modifications. In particular, the tripeptide region corresponding to Gly^{37}-Ser^{38}-Gly^{39} exhibits consistent susceptibility between different environmental changes. Strikingly, the glycine-rich cluster or GXG (X = S, T, D) tripeptide-motifs are frequently present in ATP-binding segments of various enzymes and are important for the protein-ATP interaction, enzymatic activity or protein thermal stability [39]. For the IDD_{PtkA}, whose function is still unclear, the presence of such GSG-motif may therefore represent an important regulatory feature.

The unique and unusual structural architecture of PtkA compared to other bacterial tyrosine kinases may suggest a specific role for the IDD_{PtkA}. In this context, the recently reported cross-talk between PtkA and *Mtb* endogenous eSTPK [34] is particularly interesting and foreseen to play a significant role in regulating the physiology of *Mtb*. Approximately, 7% of *Mtb* proteins undergo phosphorylation, of which 40% contain more than one phosphorylation site [35]. The high abundance of serine and threonine residues in the IDD_{PtkA} suggests potential sites for PTM. Phosphorylation of PtkA by the eukaryotic serine/threonine kinase PKA shows that residues belonging only to the IDD are phosphorylated, which may suggest a functional role of this unstructured domain. The phosphorylation of PtkA by PKA was serine specific. In contrast, the endogenous eSTPK of *Mtb* specifically phosphorylated the threonine residues of PtkA [26]. In this context, structural studies with regard to the phosphorylation of PtkA by *Mtb* endogenous serine/threonine kinases would be particularly interesting. We attempted such studies but in our hands, the eSTPK could not be stably expressed. However, endogenously expressed and purified PKA could phosphorylate IDD_{PtkA}. CSPs induced by phosphorylation of IDD_{PtkA} suggests conformational and/or local environmental changes. The largest CSPs were observed in particular for the residues located in the C-terminal part (Arg^{66}-Ala^{69}) of the IDD_{PtkA}. Strikingly, a four-fold increase in PtkA activity was observed upon phosphorylation of the IDD, thus strongly indicating a regulatory role for this domain.

In summary, we have shown that the N-terminal domain of PtkA is disordered and undergoes conformational exchange on an intermediate timescale upon varying the environmental conditions such as pH, chemical denaturant, temperature and phosphorylation. Such a dynamic property might be tightly linked to pathogenic virulence.

Acknowledgements

We thank Dr. Julia Wirmer-Bartoschek for her participation in the early stages of this work. This work was supported in part by the DKTK (German Consortium for Translational Cancer Research), FOR2509 (German Research Foundation, DFG), BioNMR (European Commission Seventh Framework Programme, grant number 261863) and iNEXT (Horizon 2020 programme of the European Union, grant number

653706). The work at BMRZ was supported by the State of Hesse.

Author contributions

HS designed and supervised the research. AN wrote the manuscript, performed and analyzed the experiments with the support of SS, HJ and HS. MH, CR, RA and CB performed measurements.

References

1 Tompa P (2002) Intrinsically unstructured proteins. *Trends Biochem Sci* **27**, 527–533.

2 Uversky VN, Oldfield CJ and Dunker AK (2005) Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *J Mol Recognit* **18**, 343–384.

3 Dyson HJ and Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* **6**, 197–208.

4 Mittag T and Forman-Kay JD (2007) Atomic-level characterization of disordered protein ensembles. *Curr Opin Struct Biol* **17**, 3–14.

5 Uversky VN (2009) Intrinsically disordered proteins and their environment: effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding. *Protein J* **28**, 305–325.

6 DeForte S and Uversky V (2016) Order, disorder, and everything in between. *Molecules* **21**, 1090.

7 Tompa P (2005) The interplay between structure and function in intrinsically unstructured proteins. *FEBS Lett* **579**, 3346–3354.

8 Rezaei-Ghaleh N, Blackledge M and Zweckstetter M (2012) Intrinsically disordered proteins: from sequence and conformational properties toward drug discovery. *ChemBioChem* **13**, 930–950.

9 Lobanov M and Galzitskaya O (2015) How common is disorder? occurrence of disordered residues in four domains of life. *Int J Mol Sci* **16**, 19490.

10 Babu MM, van der Lee R, de Groot NS and Gsponer J (2011) Intrinsically disordered proteins: regulation and disease. *Curr Opin Struct Biol* **21**, 432–440.

11 Hammoudeh DI, Follis AV, Prochownik EV and Metallo SJ (2009) Multiple independent binding sites for small-molecule inhibitors on the oncoprotein c-Myc. *J Am Chem Soc* **131**, 7390–7401.

12 Wright PE and Dyson HJ (2009) Linking folding and binding. *Curr Opin Struct Biol* **19**, 31–38.

13 Uversky VN (2013) Unusual biophysics of intrinsically disordered proteins. *Biochim Biophys Acta* **1834**, 932–951.

14 Bah A, Vernon RM, Siddiqui Z, Krzeminski M, Muhandiram R, Zhao C, Sonenberg N, Kay LE and Forman-Kay JD (2015) Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature* **519**, 106–109.

15 Kosol S, Contreras-Martos S, Cedeño C and Tompa P (2013) Structural characterization of intrinsically disordered proteins by NMR spectroscopy. *Molecules* **18**, 10802.

16 Schneider R, Huang J-R, Yao M, Communie G, Ozenne V, Mollica L, Salmon L, Ringkjobing Jensen M and Blackledge M (2012) Towards a robust description of intrinsic protein disorder using nuclear magnetic resonance spectroscopy. *Mol BioSyst* **8**, 58–68.

17 Dyson HJ and Wright PE (1998) Equilibrium NMR studies of unfolded and partially folded proteins. *Nat Struct Mol Biol* **5**(suppl), 499–503.

18 Redfield C (2004) Using nuclear magnetic resonance spectroscopy to study molten globule states of proteins. *Methods* **34**, 121–132.

19 Konrat R (2014) NMR contributions to structural dynamics studies of intrinsically disordered proteins. *J Magn Reson* **241**, 74–85.

20 Bach H, Wong D and Av-Gay Y (2009) *Mycobacterium tuberculosis* PtkA is a novel protein tyrosine kinase whose substrate is PtpA. *Biochem J* **420**, 155–162.

21 Wong D, Bach H, Sun J, Hmama Z and Av-Gay Y (2011) *Mycobacterium tuberculosis* protein tyrosine phosphatase (PtpA) excludes host vacuolar-H+–ATPase to inhibit phagosome acidification. *Proc Natl Acad Sci* **108**, 19371–19376.

22 Chow K, Ng D, Stokes R and Johnson P (1994) Protein tyrosine phosphorylation in *Mycobacterium tuberculosis*. *FEMS Microbiol Lett* **124**, 203–207.

23 Zhou P, Li W, Wong D, Xie J and Av-Gay Y (2015) Phosphorylation control of protein tyrosine phosphatase A activity in *Mycobacterium tuberculosis*. *FEBS Lett* **589**, 326–331.

24 Kusebauch U, Ortega C, Ollodart A, Rogers RS, Sherman DR, Moritz RL and Grundner C (2014) *Mycobacterium tuberculosis* supports protein tyrosine phosphorylation. *Proc Natl Acad Sci* **111**, 9265–9270.

25 Stehle T, Sreeramulu S, Löhr F, Richter C, Saxena K, Jonker HRA and Schwalbe H (2012) The apo-structure of the low molecular weight protein-tyrosine phosphatase A (MptpA) from *Mycobacterium tuberculosis* allows for better target-specific drug development. *J Biol Chem* **287**, 34569–34582.

26 Zsuzsanna Dosztányi VC (2005) Péter Tompa and István Simon. *J Mol Biol* **347**, 827–839.

27 Wishart DS, Sykes BD and Richards FM (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol* **222**, 311–333.

28 Camilloni C, De Simone A, Vranken WF and Vendruscolo M (2012) Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts. *Biochemistry* **51**, 2224–2231.
29 Marsh JA, Singh VK, Jia Z and Forman-Kay JD (2006) Sensitivity of secondary structure propensities to sequence differences between α- and γ-synuclein: implications for fibrillation. *Protein Sci* **15**, 2795–2804.
30 Goto Y, Calciano LJ and Fink AL (1990) Acid-induced folding of proteins. *Proc Natl Acad Sci* **87**, 573–577.
31 Bah A and Forman-Kay JD (2016) Modulation of intrinsically disordered protein function by post-translational modifications. *J Biol Chem* **291**, 6696–6705.
32 Uversky VN, Li J and Fink AL (2001) Evidence for a partially folded intermediate in α-synuclein fibril formation. *J Biol Chem* **276**, 10737–10744.
33 Baxter NJ and Williamson MP (1997) Temperature dependence of 1H chemical shifts in proteins. *J Biomol NMR* **9**, 359–369.
34 Zhou P, Wong D, Li W, Xie J and Av-Gay Y (2015) Phosphorylation of *Mycobacterium tuberculosis* protein tyrosine kinase A PtkA by Ser/Thr protein kinases. *Biochem Biophys Res Comm* **467**, 421–426.
35 Prisic S, Dankwa S, Schwartz D, Chou MF, Locasale JW, Kang C-M, Bemis G, Church GM, Steen H and Husson RN (2010) Extensive phosphorylation with overlapping specificity by *Mycobacterium tuberculosis* serine/threonine protein kinases. *Proc Natl Acad Sci* **107**, 7521–7526.
36 Theillet F-X, Rose HM, Liokatis S, Binolfi A, Thongwichian R, Stuiver M and Selenko P (2013) Site-specific NMR mapping and time-resolved monitoring of serine and threonine phosphorylation in reconstituted kinase reactions and mammalian cell extracts. *Nat Protoc* **8**, 1416–1432.
37 Bienkiewicz EA and Lumb KJ (1999) Random-coil chemical shifts of phosphorylated amino acids. *J Biomol NMR* **15**, 203–206.
38 Bai Y, Milne JS, Mayne L and Englander SW (1993) Primary structure effects on peptide group hydrogen exchange. *Proteins* **17**, 75–86.
39 Chen K, Wang D and Kurgan L (2015) Systematic investigation of sequence and structural motifs that recognize ATP. *Comput Biol Chem* **56**, 131–141.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Fig S1. Sequence-based analysis of PtkA.
Fig S2. Far-UV circular dichroism spectra of PtkA.
Fig S3. Analysis of the line-width and intensities of IDD_{PtkA} selected cross-peaks from 2D-(^{1}H,^{15}N)-HSQC spectra, measured at different concentrations.
Fig S4. ^{1}H,^{15}N combined CSPs and normalized peak intensities of full-length PtkA and the isolated IDD_{PtkA}.
Fig S5. Secondary chemical shifts of the IDD for IDD_{PtkA} and full-length PtkA.
Fig S6. Chemical shift analysis of IDD_{PtkA} and full-length PtkA performed by using Delta2D and SSP.
Fig S7. pH-effect on the backbone amide groups of IDD_{PtkA}.
Fig S8. GdmCl-effect on the backbone amide groups of IDD_{PtkA} at pH 7.5.
Fig S9. Temperature dependence of the amide proton chemical shift of IDD_{PtkA}.
Fig S10. Phosphorylation of PtkA by PKA.
Fig S11. MALDI spectra of IDD_{PtkA} and full-length PtkA before and after phosphorylation with PKA.
Fig S12. IDD_{PtkA} regions showing nonrandom behavior.
Fig S13. SEC chromatograms and SEC-MALS.
Table S1. Buffer used for purification of PtkA.
Appendix S1. Materials and methods.