Double Phosphorylation-Induced Structural Changes in the Signal-Receiving Domain of IκBα in Complex with NF-κB

Samira Yazdi[1], Michael Naumann[2], Matthias Stein[1*]

Short title: Double-phosphorylation events in the SRD of IκBα

Keywords: phosphoserine, degron, transcription factor inhibitor, phosphorylation, molecular dynamics, molecular signaling

[1]Max Planck Institute for Dynamics of Complex Technical Systems, Molecular Simulations and Design Group

[2]Institute of Experimental Internal Medicine, Otto-von-Guericke University Magdeburg

*Corresponding author:

Dr Matthias Stein

Molecular Simulations and Design Group

Max Planck Institute for Dynamics of Complex Technical Systems

Sandtorstrasse 1, 39106 Magdeburg, Germany.

Email: matthias.stein@mpi-magdeburg.mpg.de

Abstract

Activation of the transcription factor NF-κB requires degradation of its physiological inhibitor IκBα in order to allow nuclear translocation of NF-kB. NF-κB activity links inflammation and carcinogenesis and makes its signaling pathway an important target for therapeutic intervention. The signal-receiving N-terminal domain (SRD) of the NF-kB inhibitor IκBα harbors the sites of post-translational modifications (Ser32 and 36) directed by the IκB kinase (IKK) complex. The SRD was originally recognized to be highly disordered, but was recently shown to possess stable secondary structural elements. Identifying and characterizing the structural effects that arise as a result of phosphorylation may explain how phosphorylation regulates the IκBα-NF-κB protein complex. Therefore, the effect of post-translational mono- and double-phosphorylation of the serine residues of the SRD was analyzed. The structural modifications of the IκBα-NF-κB protein-protein complex due to mono-phosphorylation of either Ser32 or Ser36 amino acid residues or simultaneous phosphorylation were investigated by means of molecular dynamics simulations. Mono-phosphorylation at either Ser32 or Ser36 was not sufficient to induce significant structural changes in the secondary structure of the SRD of IκBα. Double-phosphorylation yielded a reduced distance between the Cα atoms of these serine residues, indicative of a structural change. Only this two-fold phosphorylation induced the extended conformation of the degron motif which renders it accessible by the E3 ligase. In summary, these results provide insight into the conformational changes induced in IκBα proteins upon phosphorylation that are vital to their signaling dynamics and enable us to propose a model for the phosphorylation of the SRD.

Introduction

The cytoplasmic inhibitor IκBα (nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha) of NF-κB ('nuclear factor kappa-light-chain-enhancer' of activated B-cells) was discovered first in 1988 by Baeuerle and Baltimore [1]. The physiological inhibitor of NF-κB, IκBα, is responsible for a systems control of the process of NF-κB signaling. Activation of the NF-κB pathway is associated with inflammation, stress and developmental processes and immune functions [2,3], and its improper regulation results in a variety of disorders including a wide range of cancers, neurodegenerative diseases, asthma, arthritis, inflammatory bowel disease, and multiple sclerosis [4,5]. The most common inhibitors of NF-κB dimers, in addition to IκBα, are IκBβ, IκBε, and BCL-3, all of which block the nuclear localization and transcriptional activity of NF-κB [6,7]. All of these inhibitory proteins contain multiples of a repeat structure known as the ankyrin repeat, a canonical 33 amino acid sequence motif that enables protein-protein interactions [6].

(Scheme 1)

The series of key events, which trigger NF-κB activation, have particularly been well-characterized. The 'canonical' activation pathway of NF-κB [8], which requests degradation of IκBα (see Scheme 1), is initiated by the critical step of phosphorylation [9] by the protein IκB kinase (IKK) [10,11] in response to distinct stimuli such as tumor necrosis factor (TNF) and interleukin 1 (IL-1), T and B cell mitogens, bacterial lipopolysaccharide (LPS), viruses, and double-stranded (ds) RNA [12-14]. The DNA binding NF-κB complex is composed of a heterodimer (RelA and p50) in which RelA carries the transactivation domain (see **Figure 1**

A). A second NF-κB activation pathway has been termed the 'non-canonical' pathway – this pathway entails that processing of the p100 subunit results into the N-terminal DNA-binding protein (p52) and nuclear translocation of the processed p52 and the associated RelB protein.[15]

In the 'canonical' activation pathway, phosphorylation subsequently marks the signal-induced recognition of IκBα at sites Ser32 and Ser36 by the F-box/WD40 protein β-TrCP [16], which in turn prompts the ubiquitination of IκBα N-terminal residues Lys21 or Lys22 by the E3 ubiquitin ligase, SKp1–Cullin–F-box (SCF) [17-20]. Although phosphorylation of IκBα precedes its dissociation from NF-κB, it is not sufficient and requires additional ubiquitination to enable its dissociation and subsequent degradation [21]. Other data suggest that timely and efficient degradation of ubiquitinated IκBα takes place directly when still associated with RelA [22]. Further, timely and efficient liberation of RelA from ubiquitinated IκBα and RelA nuclear translocation, essentially depends on the presence of functional p97/VCP.

(Figure 1)

Amino acid substitutions of the two adjacent potential ubiquitination sites in the N-terminus of IκBα (Lys21 and Lys22) almost completely block the rapid, signal-induced degradation of the mutant protein, while they do not interfere with kinase-induced phosphorylation [19,23].

Following these events, IκBα is degraded by the 26S proteasome as a final step in this signaling cascade potentiating the translocation of NF-κB into the nucleus. The ubiquitin E3 ligase only recognizes a double-phosphorylated IκBα through a well-known recognition sequence, the degron motif [24], fairly conserved among IκB inhibitors and the signal-transducing protein β-catenin (**Figure 1 B**) [25]. The degron sequence is contained within the

signal receiving domain (SRD) in the N-terminus of IκBα, adjacent to the ankyrin repeat region (**Figure 1 C**). Composed of residues DSGϕXS, with ϕ denoting a hydrophobic residue and X any residue, the degron motif contains the vital phosphorylation sites, namely serine residues Ser32 and Ser36. Mutation of either residue disrupted polyubiquitination of IκBα and abolished the stimulus-induced degradation of IκBα, thereby requiring the strict phosphorylation of both serine residues as a recognition and initial step in NF-κB activation[21,26-30].

In our approach to unravel the molecular basis of NF-κB signaling by multiscale molecular simulations, we initially refined the previously annotated 'disordered' SRD of IκBα to be structured with additional helical secondary structure elements based on a consensus of multiple secondary structure element predictions and their stability in long molecular dynamics (MD) simulation for free IκBα and when in complex with NF-κB [31]. We constructed a structural model of full length IκBα by modeling the missing N-terminal segment in the only in part available crystal structure of the IκBα/NF-κB complex (**Figure 1 A**)[31]. By performing microsecond MD simulations, we investigated the dynamic properties of the unphosphorylated IκBα both in its free and NF-κB-bound states. These findings provide the starting platform for investigating and understanding the mechanism of phosphorylation in IκBα by IKK and the effect of single and double phosphorylation by IKK. To paint a complete picture of how phosphorylation guides and regulates protein function, the rub lies in understanding the molecular basis of this phenomenon. With phosphorylation being the most well-studied posttranslational modification, computational simulation methods on the atomistic scale have proven valuable in offering explanations of the structural basis of protein regulation by phosphorylation in addition to complementing experimental studies[32-34].

The crystal structure of the β-TrCP/β-catenin complex has revealed the basis of substrate recognition by the WD40 domain in β-TrCP and shed light on the intimate interaction between degron motif containing substrates and the SCF ubiquitin ligase [35]. Of similar interest would be the interactions between β-TrCP and the phosphorylated IκBα; however prior to this, the subject of interest is the structural rearrangement triggered by phosphorylation of the SRD in the ankyrin repeat protein which drives the recognition mechanism by the SCF complex. To this end, we have in this study examined the role of mono- and double-phosphorylation by probing into the structural effects that it induces on IκBα bound to its partner, NF-κB. By applying multiple MD simulations of 500 ns each, phosphorylation of Ser32 and Ser36 of IκBα in complex with NF-κB, both in their double- and mono-states of phosphorylation were studied. Our study identified a partially more stable region locally by the site of phosphorylation in the double-phosphorylated IκBα. Moreover, double-phosphorylation induced a rather extended conformation in the vicinity of the phosphorylation site as compared to a more bent shape observed in the unphosphorylated state. The mono-phosphorylated states each fell into an intermediary state between the unphosphorylated and the double-phosphorylated states, partially initiating this structural transition. In addition, distinct variations between pSer32 and pSer36 were observed. pSer32 emerged as the more solvent exposed residue, slightly burying Ser36 and inducing formation of a new hydrogen bond pattern stabilizing the N-terminal tail and the region closest to pSer36. Our findings also revealed a more pronounced electrostatic effect upon double-phosphorylation, which induces structural rearrangements that change the surface charge potential and create a greater acidic environment around the phosphorylation site. Overall, our findings explain the prerequisite for double-phosphorylation on a detailed molecular level and offer an insight into the structural rearrangements that take place, hence laying the ground work into future studies of IκBα/β-TrCP recognition and binding mode.

Materials and Methods

Structural model of IκBα/NF-κB

The model of the full-length IκBα/NF-κB complex was obtained as described in a previous study [31] containing the complete signal receiving domain (SRD; residues 1-67) of IκBα and the ankyrin repeat domain (ARD; 68-280), thereby providing a suitable working model to investigate the phosphorylation in the SRD region of IκBα. A model of a double-phosphorylated simulation system is presented in **Figure 1 A**.

Molecular simulation setup

The well-equilibrated structural model of the unphosphorylated IκBα/NF-κB complex [31] was used as an initial structure for the simulations. To investigate the role of phosphorylation, four distinct system setups were constructed: an unphosphorylated (nIκBα) system, two mono-phosphorylated complexes where either of serines 32 (p32IκBα) or 36 (p36IκBα) were post-translationally modified into phosphoserines, and a double-phosphorylated system (ppIκBα) in which serine residues 32 and 36 were both transformed into phosphoserines simultaneously. The phosphoserine mutations were introduced with PyMol[36]. The simulations were carried out with Gromacs 4.5 [37] with the modified GROMOS96 43a1p force field [38] containing the parameters for phosphoserine. This force field was shown to give a balanced description of secondary structural elements without any bias or preference[39,40]. A time step of 2 fs was used. The LINCS algorithm [41] was applied for constraining bond lengths. Electrostatic interactions were calculated with the Particle-Mesh Ewald algorithm at every step [42]. A 1.0 nm cutoff was used both for electrostatics and van der Waals interactions, with neighborlists updated every 10 steps. The simulations were performed at constant pressure of 1.0 bar with Parrinello-Rahman pressure coupling [43] and the isotropic pressure

scaling, time constant of 1.0 ps, and a system compressibility of 4.5e-5 bar^{-1}. The temperature of the system was coupled to 300 K using the velocity-rescaling algorithm [44]. The nIκBα system containing 6945 protein atoms was solvated with 31,361 SPC water molecules, 218 out of which were replaced with 122 sodium and 96 chloride ions to neutralize the total net system charge and obtain a physiological salt concentration of 0.15 M. The total number of atoms in the wild-type system reached 100,226 all of which were placed in a rectangular cell of an approximate system size of 9.4 x 8.0 x 13.7 nm^3. For the post-translationally modified systems, sodium ions were replaced with corresponding number of water molecules to maintain a neutralized system charge depending on the phosphorylation state of the systems. The systems were energy minimized with steepest descent until the maximum force reached < 100 kJ/mol/nm and were subsequently equilibrated for 100 ps while keeping the protein position restrained (F_c = 1000 kJ/mol/nm^2). Eventually, all position restraints were removed and the production runs were performed for 500 ns each. Three independent replicates of each system were simulated, each starting with different initial velocities, amounting to a total simulation time of 6 μs, thus allowing a thorough exploration of the extensive conformational space.

Results

In order to study the effects of phosphorylation of Ser32 and Ser36 located on the SRD of IκBα, we carried out MD simulations of the unphosphorylated and the double-phosphorylated IκBα in complex with the transcription factor NF-κB. To further characterize the solitary roles of mono-phosphorylation of Ser32 and Ser36, two additional mono-phosphorylated systems were studied and compared with the unphosphorylated and the double-phosphorylated IκBα/NF-κB complex. The reported results are based on an average

finding of the three independent replicate runs of each system, which were simulated for 500 ns each.

Local divergence in structural stability promoted by double-phosphorylation

The degree of flexibility of the protein can be evaluated by measuring the backbone atom root mean square fluctuation (RMSF) around the initial structure. The stretch of residues, IκBα$^{28\text{-}40}$, encompassing the sites of phosphorylation Ser32 and Ser36 as shown in **Figure 2 A**, reveal a difference in flexibility between the unphosphorylated and the double-phosphorylated system. For each system, the trajectories were individually averaged and we give the average RMSF with the standard error. For the unphosphorylated system, IκBα$^{28\text{-}40}$ displays an average RMSF of 1.35 ± 0.2 Å, in contrast to a lower RMSF of 1.08 ± 0.1 Å for the double-phosphorylated system (**Figure 2 B**).

(Figure 2)

In the mono-phosphorylated systems, a lower RMSF is observed between that of the unphosphorylated and double-phosphorylated systems (p32 1.21 ± 0.2 Å and p36 1.07 ± 0.1 Å). The results for the individual trajectories can be found in the Supporting Information **(Table S1 and Figure S1)**. However, the RMSF curves for the mono-phosphorylated systems remain rather varied across the segment, resembling much that of the unphosphorylated system as evident from **Figure 2 C**. In contrast, double phosphorylation results in a fairly stable segment as revealed by the RMSF curve. Compared to the unphosphorylated system, the RMSFs of the residues preceding and succeeding the phosphorylation sites up to residue 38 in the double-phosphorylated system remain relatively lower. The general observations in the same stretch of residues are supported by the root mean square deviation (RMSD) of the backbone atoms of this region (**Figure 2 D**). The

average RMSD of this segment settles at 2.5 ± 0.6 Å for the unphosphorylated system, opposed to a rather lower RMSD of 1.7 ± 0.2 Å in the event of double phosphorylation. Similar to the trend observed for the atomic fluctuations, mono phosphorylation causes the average RMSDs to fall between that of the unphosphorylated and double-phosphorylated systems. The details for each of the trajectories are available in the Supporting Information (**Table S2 and Figure S1 and S2**). The RMSD against time plot (**Figure 2 E**) reveals that the difference in RMSD of the unphosphorylated and double-phosphorylated systems is detectable from very early on in the simulation and evident for the entire simulation. In general, double phosphorylation leads to a structurally more stable region surrounding the sites of phosphorylation, while mono phosphorylation causes partially stable structures with pSer36 modification having a larger effect compared to pSer32 modification. In order to evaluate the global stability of the protein-protein complex before and after phosphorylation, the RMSD of all backbone atoms of the IκBα/NF-κB complex was calculated for all systems (**Figure S1** shows the RMSD plots for all three replicate runs for all three phosphorylation states). Overall, the IκBα/NF-κB complex simulations remain stable throughout all states.

Double-phosphorylation induces a local N-terminal conformational change in the SRD

To further characterize the induced structural variation due to the two-fold phosphorylation, we measured the Cα-Cα distance between Ser32 and Ser36 plotted against time (**Figure 3 A**). By mapping the time evolution of the main-chain Cα distances between these residues, we investigate the structural effect of mono-phosphorylated pSer32/nSer36, nSer32/pSer36 and double-phosphorylated pSer32/pSer36 in detail, and compare with the unphosphorylated nSer32/nSer36 system.

(Figure 3)

A difference in the Ser32$^{C\alpha}$-Ser36$^{C\alpha}$ distances could be observed for the unphosphorylated and the double-phosphorylated complexes. In the double phosphorylated state, the Cα-Cα distance increases gradually in the first 30 ns of the simulation, after which it reaches a distance of 13 Å and maintains it throughout the rest of the simulation. In contrast, the Cα-Cα distance in the unphosphorylated state drops sharply down to 9 Å in about 30 ns and sustains this distance at about 9-10 Å for the remaining of the 500 ns simulation duration averaged over three independent simulation runs. In both of the mono-phosphorylated systems, the Cα-Cα distance fluctuates between 9-12 Å in the initial 300 ns of the simulation, however, into the final 200 ns both systems keep a rather constant distance at approximately 12 Å. This shows that the two-fold phosphorylation has a larger effect than mono-phosphorylation of either Ser32 or Ser36 residues but that the structural effect is not additive (Cα-Cα distance plots for each of the trajectories can be found in **Figure S3**). These observations are reflected clearly in **Figure 3 B** that displays the average distances for the different systems. With an average Ser32$^{C\alpha}$-Ser36$^{C\alpha}$ distance of 9.4 Å and 12.8 Å in the unphosphorylated and double-phosphorylated states, respectively, double phosphorylation leads to an increase of 3.4 Å in this Cα-Cα distance between the two sites of phosphorylation. This variation in the Ser32$^{C\alpha}$-Ser36$^{C\alpha}$ distance between the unphosphorylated and the double-phosphorylated states is inherent in the structural conformation of the phosphorylation region. The representative conformations of a stretch of residues 31 to 37 in the unphosphorylated (**Figure 3 C**) and the double-phosphorylated (**Figure 3 D**) complexes, clearly show a region with a defined bend in between the sites of phosphorylation in the unphosphorylated system, and a slightly more extended structure is revealed upon two-fold phosphorylation.

Variation in solvent exposure in Ser32 and Ser36

The impact of phosphorylation on sites Ser32 and Ser36 can be assessed by calculating the relative solvent accessible surface area (SASA) of the individual serine residues.

(Figure 4)

As evident from **Figure 4 A**, in the unphosphorylated complex, both Ser32 and Ser36 remain surface exposed (staying above a threshold of 20 %), with Ser36 displaying a marginal fluctuation around a relative SASA of 40%.[31] Intriguing are also the structural observations of the effect of phosphorylation relative to the unphosphorylated state. Mono-phosphorylation of pSer32 sets in motion structural rearrangements, which in turn induce a partial burial of Ser36 as shown in **Figure 4 C**. However, the mono-phosphorylation of pSer36 triggers first the increase of the relative SASA of pSer32 as compared to the unphosphorylated structure, and second, it allows a higher surface exposure of pSer36, even higher than the pSer36 relative SASA in the double-phosphorylated state (**Figure 4 B**). In clear contrast to the unphosphorylated state, double-phosphorylation leads to a striking difference in the relative SASA between pSer32 and pSer36 (**Figure 4 D**). Upon phosphorylation, pSer36 maintains its relative SASA of 40 %, suggesting a partial burial of this residue. However, pSer32 maintains a highly exposed configuration with a relative SASA of 70 % during a larger part of the simulation, allowing the potential for a wide range of interactions (**Figure S4 A**). These observations are suggestive of the distinctive roles that pSer32 and pSer36 could play upon phosphorylation: a rather exposed pSer32 which extends beyond the surface of the protein to act as an anchoring point and engage in inter-protein interactions, and a comparatively less exposed pSer36 that could contribute to key-specific intra-protein interactions.

Variation in solvent exposure, due to phosphorylation in the vicinity of the sites of phosphorylation, has been mapped for individual residues in segment $I\kappa B\alpha^{1-40}$ as displayed in **Figure S4 B**. The relative SASA of the residues located in between Ser32 and Ser36

remains unchanged except for Leu34. With its hydrophobic nature conserved in the degron motif and known to make interactions with β–TrCP[35], Leu34 becomes partly surface exposed in the double-phosphorylated state from having been rather buried in its natural state. Further, the ubiquitination site Lys21, which is relatively buried in its natural state, increases its surface exposure upon double phosphorylation and thus its accessibility for ubiquitination. The absolute SASA for the four states of phosphorylation can be found in the Supplementary Information (**Figures S5 and S6**).

Double-phosphorylation stabilizes region by novel hydrogen bond interactions

The conformational differences caused by phosphorylation may be accounted for by a new hydrogen bond formation pattern in the double-phosphorylated complex. Based on our simulations, out of the two phosphorylation sites, it is in particular pSer36 that contributes to establishing newly formed interactions.

(Figure 5)

Residues Met1 and Gln3 located at the tip of the N-terminal shift away from pSer32 and closer to pSer36 and form unique hydrogen bonds: The backbone amide of Met1 interacts with the phosphate group of pSer36, whereas both the backbone and side chain amides of Gln3 form hydrogen bonds with the backbone carbonyl group of pSer36; these shifts together with other local conformational changes lead to the stabilization of the N-terminal tail and concomitant other regional changes in the structure of IκBα (**Figure 5 A**). Additional residues that form additional hydrogen bonds in the vicinity of the phosphorylation site are Asp35/Gln44 (**Figure 5 B**) and Met37/Glu43 (**Figure 5 C**). The backbone carbonyl of Asp35 engages in a hydrogen bond interaction with the side chain amide of Gln44, and the backbone amide of Met37 forms a similar interaction with the side chain carboxyl group of Glu43. Interestingly, Asp35 and Met37 both enclose the pSer36 phosphorylation site, and the

adjoining residues Glu43 and Gln44 are localized on the third α-helix in the SRD of IκBα. The relevant hydrogen bonds formed between donor and acceptor are given in **Table 1** together with their time of persistence during the simulation runs.

Hydrogen bond formation details for different donor and acceptor atoms are given in the Supplementary Information (**Table S3**).

(Table 1)

The formation of additional interactions present only in the double-phosphorylated and absent in the unphosphorylated states is in agreement with experimental findings from Pons[45] who identified intramolecular hydrogen bonds in a truncated 24 peptide segment IκBα$^{21-44}$ bound to *β*-TrCP. In the complex, intramolecular hydrogen bonds involving residues Arg24-Asp31 and Met37-Gln44 stabilized the peptide conformation (see below). These findings emphasize the role that phosphorylation possesses in establishing intramolecular stability and order in the region, as reflected previously in the lower RMSF exhibited by the double-phosphorylated state. Specifically, it appears that phosphorylation of Ser36 is the principal contributor to this effect.

Electrostatic effects

Part of the effects of phosphorylation may be explained by differences in the electrostatic potential surrounding the site of phosphorylation. The degron motif in IκBα$^{31-36}$ holds the sequence of residues DSGLDS, which encompasses the integral site of phosphorylation Ser32 and Ser36, each of which are preceded by an acidic residue, aspartic acid. The electrostatic potential of this region is illustrated in **Figure 6**, with red colored surface correlated with potentials of -10 kT/e and blue colored surface correlating with potentials of +10 kT/e calculated using APBS.[46]

(Figure 6)

In the unphosphorylated state (**Figure 6 A**), the negative potential patches at or near the protein surface surrounding the sites of phosphorylation are fostered by the presence of strong acidic residues, namely Asp27, Asp28, Asp31 in the α-helical region preceding Ser32 and residues Asp39, Glu40, Glu41, Glu43 that are localized on the loop segment following Ser36 [31]. In the double-phosphorylated state, the previously negatively charged patch grows even stronger covering a more extended part of the protein surface (**Figure 6 D**). The introduced negative charges upon phosphorylation, jointly with induced structural rearrangements alter the distribution of surface charge potential in a visible manner creating an even more pronounced negatively charged protein surface (see also **Figure S7**). The mono-phosphorylated states, as displayed by **Figure 6 B** and **C**, do not exhibit the same electrostatic effects as their double-phosphorylated counterpart. Since double-phosphorylation is a prerequisite for β-TrCP binding, the electrostatic surface potential of the binding area of β-TrCP is shown (**Figure 6 E**). The surface potential of the top narrow part of the channel being the binding surface reveals an extensive positive blue colored patch indicative of a basic environment. The electrostatic complementarity of double-phosphorylated IκBα and β-TrCP protein surfaces may be the critical recognition mechanism to initiate the formation of the tertiary protein complex.

Discussion

Phosphorylation is the most prevalent post-translational mechanism regulating protein function throughout the cell. Protein kinases carry out phosphorylation by modifying existing serine, threonine or tyrosine side chains with the addition of a phosphate group. At physiologically relevant pH, a phosphate group usually carries a -2 charge. Introducing such a negative charge often leads to electrostatic perturbations directly affecting protein energy

landscapes, which exercise control over protein-protein interactions and conformational dynamics [47]. As with many other cellular functions, it is challenging to derive an atomic-level understanding of how phosphorylation alters protein structure and function. Computational methods, such as all atomistic MD simulations, are a major contributor in filling this knowledge-gap and throwing light on the structural changes upon protein phosphorylation (as reviewed in [47]).

Many experimental studies on the importance of IκBα phosphorylation have been done in the past [21,26-30]. Yet little is known about the molecular details which render this post-translational modification indispensable for recognition by the SCF complex. With MD simulations, we were able to characterize local structural aspects induced by phosphorylation of Ser32 and Ser36. Although no striking differences were observed in secondary structure elements upon phosphorylation, aspects of local stability, in the progression of disorder to order surrounding the sites of phosphorylation were evidenced. Double-phosphorylation leading to the uniform ordering of the segment IκBα$^{28-40}$ was brought forth by the stabilization of residues preceding the initial phosphorylation site and succeeding the second phosphorylation site, namely residues 28-38. Moreover, increased N-terminal stabilization induced by forming hydrogen bond interactions with the phosphorylation site was another direct effect of double-phosphorylation. In contrast, the unphosphorylated IκBα experienced a more variable structural stability across this region. Post-translational phosphorylation has often previously been associated with sparking both ordering and disordering in proteins. For the degron of IκBα in complex with NF-κB we observe an ordering effect upon phosphorylation which was also observed for MD simulations of myosin.[48]

The effect of phosphorylation in proteins can be manifold. This PTM can induce conformational changes, promote order–disorder transitions and modulate recognition via electrostatic interactions with binding partners. Groban et al. [34]described the use of MD

simulations to map loop or helices induced by phosphorylation. They concluded that the induced changes are local in nature, limited to relatively modest conformational changes and not, e.g., more drastic order–disorder transitions.

A Monte Carlo/Stochastic Dynamics simulations study revealed how electrostatic interactions navigated the stabilization of a helical conformation at the N-terminus upon phosphorylation in model peptides consisting of a serine in the N-terminus followed by nine alanines [49]. Likewise, phosphorylation of the phenylalanine hydroxylase (hPAH) lead to an increase in stability of the N-terminal tail through local conformational changes as a result of electrostatic interactions[50]. Car-Parrinello simulations showed that phosphorylation of the CREB-CBP complex resulted in an increased stability of the complex by forming a new hydrogen bond interaction[51]. In another study employing molecular mechanics–based methods on kinases and prokaryotic response regulators, local conformational changes specific to areas in close proximity to the phosphorylated amino acid were observed [34]. Structural transitions between order and disorder induced by phosphorylation have also been mapped by previous experimental work. An order-to-disorder transition of a helix containing a phosphorylated serine was documented for the oncoprotein 18/stathmin (Op18)[52], in contrast to a disorder-to-order adaptation of the protein kinase B/Akt which experiences disruption of the αC-helix causing a global restructuring of the protein upon phosphorylation.[53]

Our unphosphorylated state simulations indicate a weakly bent shape between the two serine residues. However, an impact of double-phosphorylation was the structural reconfiguration of the segment connecting the phosphoserines resulting in an extended conformation and verified by a larger $Ser32^{C\alpha}$-$Ser36^{C\alpha}$ distance. A possible explanation is the electrostatic repulsion by the two negatively charged phosphorylated side chain serine residues in the stretch **$DpS^{32}GLDpS^{36}$** which are only 3 amino acids apart. The presence of additional

negatively charged aspartate residues D^{31} and D^{35} immediately preceding the serine residues also may contribute to this effect.

Pons *et al.* [45] investigated a truncated 24 amino acid peptide ppIκBα$^{21-44}$ by NMR in free solution and in complex with *β*-TrCP. The segment was doubly phosphorylated at Ser32 and Ser36 to mimic the SRD of IκBα. In solution they observed that the free peptide behaves like a random coil with repulsion between the two phosphate groups and the tendency to form a large bend. This situation is non-physiological since ppIκBα does not exist in solution and when in complex with NF-κB would immediately attract the E3 ubiquitin ligase SCF and cannot be compared to our simulations. In the presence of *β*-TrCP, however, 24 ppIκBα adopts a well-defined three-dimensional structure consisting of a central bend from residues 29 to 37. This is in agreement with our MD simulations which showed the decrease of atomic fluctuations and RMSD in the doubly phosphorylated state and formation of a more ordered state. Our calculated phosphate group distance between pSer32 and pSer36 is 15.27 ± 0.27 Å which is in excellent agreement with the 15.5 Å observed in the NMR structures. [45]

A distinct pattern of conduct of the phosphorylated serine residues, revealed by the simulations, is indicative of the individual roles played by each serine residue. We propose that phosphorylation is commenced at site Ser36, which in turn increases the solvent surface exposure of Ser32 increasing its accessibility to the kinase IKK, although partially burying pSer36 (**Scheme 2**).

(Scheme 2)

This relative partial burial of pSer36 leads to a new hydrogen bond network stabilizing the N-terminal segment of IκBα and the region proximal to the pSer36 phosphorylation site. With pSer36 contributing to local structural stability of the protein, pSer32 with a relatively large SASA is able to interact freely with β-TrCP. This view is supported by the resolved crystal

structure of the bound β-TrCP to β-catenin showing how pSer33 (homologous to pSer32 in IκBα) is the residue that makes the largest number of contacts with β-TrCP, that is residues Tyr271, Ser309, Ser325 and Arg285, while pSer37 (homologous to pSer36 in IκBα) forms comparatively fewer interactions with residues Ser448, Gly432 and Arg431 (**Figure 6 E**) [35]. These residues are located on the rim at opposite sides of the channel. Moreover, molecular docking data of a phosphorylated 11 amino acid IκBα peptide bound to β-TrCP demonstrated that pSer32 establishes the same interactions as β-catenin [54].

Long-range electrostatic effects are a major determinant of protein-protein recognition and association. The electrostatic complementarity of protein surfaces was identified to be a major regulator of protein-protein complex formation, see for example [55-57]. Post-translational introduction of phosphate groups carrying a double negative charge by IKK enhances the negative electrostatic potential in the double-phosphorylated state of IκBα, an effect remarkably different from the unphosphorylated and the mono-phosphorylated states. The elicited negatively charged protein environment adjacent to the phosphorylation sites is also of critical importance in the recognition by the β-TrCP SCF complex.

Conclusion

We performed full atomistic MD simulations for a total of 6 μs on unphosphorylated, the two mono-phosphorylated and a double-phosphorylated IκBα protein in complex with NF-κB. We were able to find support for a sequential two-step phosphorylation by IKK. In the unphosphorylated nIκBα, Ser36 displays a larger solvent accessibility and probably is the first site of phosphorylation by IKK. pSer36 leads to a stabilization of the conformation of the N-terminal region and renders Ser32 to interact with the kinase and thus become post-translationally modifiable. The two-fold phosphorylation induces a disorder-to-order

transition of the degron and exposes pSer32 and pSer36 to be recognizable by binding partners. In the double-phosphorylated form, ppIκBα, pSer32 is the major site of interaction with β-TrCP.

This is in agreement with experimental studies, which showed that when either Ser32 or Ser36 of IκBα was mutated, the protein was not phosphorylated and did not undergo degradation, and NF-κB could not be activated.[27] In addition, the Ser32A/Ser36A double mutant of IκBα was neither phosphorylated nor degraded in response to signal induction and failed to undergo inducible ubiquitination *in vivo*. We can conclude that pSer32/pSer36 double-phosphorylation does not directly lead to a dissociation of the IκBα/NF-κB complex but initiates a conformational change of the degron to enable binding to β-TrCP which then initiates ubiquitination of the IκBα N-terminal residues Lys21 or Lys22 by the E3 ubiquitin ligase, SCF. Thus, signal-induced activation of NF-κB involves phosphorylation-dependent ubiquitination of IκBα, which subsequently targets the protein for rapid degradation by the proteasome and releases NF-κB for translocation to the nucleus, see for example [58]. This will be the topic of further investigations of the molecular systems biology in the signal transduction pathway of IκBα signaling.

Acknowledgements

This work was financially supported by The Max Planck Society for the Advancement of Science, the Excellence Initiative "Research Center for Dynamic Systems: Systems Engineering" by the state of Saxony-Anhalt and the "European Regional Development Fund "Europäischer Fonds für regionale Entwicklung (ERDF)".

References

1. Baeuerle PA, Baltimore D. I kappa B: a specific inhibitor of the NF-kappa B transcription factor. Science 1988;242(4878):540-546.
2. Baeuerle PA, Henkel T. FUNCTION AMD ACTIVATION OF NF-KAPPA-B IN THE IMMUNE-SYSTEM. Annual Review of Immunology 1994;12:141-179.
3. Baldwin AS. The NF-kappa B and I kappa B proteins: New discoveries and insights. Annual Review of Immunology 1996;14:649-683.
4. Baeuerle PA, Baltimore D. NF-kappa B: ten years after. Cell 1996;87(1):13-20.
5. Kumar A, Takada Y, Boriek AM, Aggarwal BB. Nuclear factor-kappaB: its role in health and disease. J Mol Med (Berl) 2004;82(7):434-448.
6. Verma IM, Stevenson JK, Schwarz EM, Van Antwerp D, Miyamoto S. Rel/NF-kappa B/I kappa B family: intimate tales of association and dissociation. Genes Dev 1995;9(22):2723-2735.
7. Basak S, Kim H, Kearns JD, Tergaonkar V, O'Dea E, Werner SL, Benedict CA, Ware CF, Ghosh G, Verma IM, Hoffmann A. A fourth IkappaB protein within the NF-kappaB signaling module. Cell 2007;128(2):369-381.
8. Beg AA, Baldwin AS, Jr. The I kappa B proteins: multifunctional regulators of Rel/NF-kappa B transcription factors. Genes Dev 1993;7(11):2064-2070.
9. Ghosh S, Baltimore D. Activation in vitro of NF-kappa B by phosphorylation of its inhibitor I kappa B. Nature 1990;344(6267):678-682.
10. Chen ZJ, Parent L, Maniatis T. Site-specific phosphorylation of I kappa B alpha by a novel ubiquitination-dependent protein kinase activity. Cell 1996;84(6):853-862.
11. DiDonato JA, Hayakawa M, Rothwarf DM, Zandi E, Karin M. A cytokine-responsive IkappaB kinase that activates the transcription factor NF-kappaB. Nature 1997;388(6642):548-554.
12. Siebenlist U, Franzoso G, Brown K. Structure, regulation and function of NF-kappa B. Annu Rev Cell Biol 1994;10:405-455.
13. Ghosh S, May MJ, Kopp EB. NF-kappa B and Rel proteins: evolutionarily conserved mediators of immune responses. Annu Rev Immunol 1998;16:225-260.
14. Rothwarf DM, Karin M. The NF-kappa B activation pathway: a paradigm in information transfer from membrane to nucleus. Sci STKE 1999;1999(5):RE1.
15. Bonizzi G, Karin M. The two NF-kappaB activation pathways and their role in innate and adaptive immunity. Trends Immunol 2004;25(6):280-288.
16. Yaron A, Hatzubai A, Davis M, Lavon I, Amit S, Manning AM, Andersen JS, Mann M, Mercurio F, Ben-Neriah Y. Identification of the receptor component of the IkappaBalpha-ubiquitin ligase. Nature 1998;396(6711):590-594.
17. Palombella VJ, Rando OJ, Goldberg AL, Maniatis T. The ubiquitin-proteasome pathway is required for processing the NF-kappa B1 precursor protein and the activation of NF-kappa B. Cell 1994;78(5):773-785.
18. Chen Z, Hagler J, Palombella VJ, Melandri F, Scherer D, Ballard D, Maniatis T. Signal-induced site-specific phosphorylation targets I kappa B alpha to the ubiquitin-proteasome pathway. Genes Dev 1995;9(13):1586-1597.
19. Baldi L, Brown K, Franzoso G, Siebenlist U. Critical role for lysines 21 and 22 in signal-induced, ubiquitin-mediated proteolysis of I kappa B-alpha. J Biol Chem 1996;271(1):376-379.
20. Karin M, Ben-Neriah Y. Phosphorylation meets ubiquitination: The control of NF-kappa B activity. Annual Review of Immunology 2000;18:621-663.
21. DiDonato JA, Mercurio F, Karin M. Phosphorylation of I kappa B alpha precedes but is not sufficient for its dissociation from NF-kappa B. Mol Cell Biol 1995;15(3):1302-1311.
22. Schweitzer K, Pralow A, Naumann M. p97/VCP promotes Cullin-RING-ubiquitin-ligase/proteasome-dependent degradation of I κ B α and the preceding liberation of RelA from ubiquitinated I κ B α . Journal of Cellular and Molecular Medicine 2016;20(1):58-70.
23. Winston JT, Strack P, Beer-Romero P, Chu CY, Elledge SJ, Harper JW. The SCFbeta-TRCP-ubiquitin ligase complex associates specifically with phosphorylated destruction motifs in

IkappaBalpha and beta-catenin and stimulates IkappaBalpha ubiquitination in vitro. Genes Dev 1999;13(3):270-283.
24. Yaron A, Gonen H, Alkalay I, Hatzubai A, Jung S, Beyth S, Mercurio F, Manning AM, Ciechanover A, Ben-Neriah Y. Inhibition of NF-kappa-B cellular function via specific targeting of the I-kappa-B-ubiquitin ligase. EMBO J 1997;16(21):6486-6494.
25. Aberle H, Bauer A, Stappert J, Kispert A, Kemler R. beta-catenin is a target for the ubiquitin-proteasome pathway. EMBO J 1997;16(13):3797-3804.
26. Brockman JA, Scherer DC, McKinsey TA, Hall SM, Qi X, Lee WY, Ballard DW. Coupling of a signal response domain in I kappa B alpha to multiple pathways for NF-kappa B activation. Mol Cell Biol 1995;15(5):2809-2818.
27. Brown K, Gerstberger S, Carlson L, Franzoso G, Siebenlist U. Control of I kappa B-alpha proteolysis by site-specific, signal-induced phosphorylation. Science 1995;267(5203):1485-1488.
28. DiDonato J, Mercurio F, Rosette C, Wu-Li J, Suyang H, Ghosh S, Karin M. Mapping of the inducible IkappaB phosphorylation sites that signal its ubiquitination and degradation. Mol Cell Biol 1996;16(4):1295-1304.
29. Traenckner EB, Pahl HL, Henkel T, Schmidt KN, Wilk S, Baeuerle PA. Phosphorylation of human I kappa B-alpha on serines 32 and 36 controls I kappa B-alpha proteolysis and NF-kappa B activation in response to diverse stimuli. EMBO J 1995;14(12):2876-2883.
30. Whiteside ST, Ernst MK, LeBail O, Laurent-Winter C, Rice N, Israel A. N- and C-terminal sequences control degradation of MAD3/I kappa B alpha in response to inducers of NF-kappa B activity. Mol Cell Biol 1995;15(10):5339-5345.
31. Yazdi S, Durdagi S, Naumann M, Stein M. Structural modeling of the N-terminal signal-receiving domain of IkappaBalpha. Front Mol Biosci 2015;2:32.
32. Perkins ND. Post-translational modifications regulating the activity and function of the nuclear factor kappa B pathway. Oncogene 2006;25(51):6717-6730.
33. Scheidereit C. I kappa B kinase complexes: gateways to NF-kappa B activation and transcription. Oncogene 2006;25(51):6685-6705.
34. Groban ES, Narayanan A, Jacobson MP. Conformational changes in protein loops and helices induced by post-translational phosphorylation. Plos Computational Biology 2006;2(4):238-250.
35. Wu G, Xu G, Schulman BA, Jeffrey PD, Harper JW, Pavletich NP. Structure of a beta-TrCP1-Skp1-beta-catenin complex: destruction motif binding and lysine specificity of the SCF(beta-TrCP1) ubiquitin ligase. Mol Cell 2003;11(6):1445-1456.
36. Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.3r1. 2010.
37. Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 2008;4(3):435-447.
38. Scott WRP, Hunenberger PH, Tironi IG, Mark AE, Billeter SR, Fennen J, Torda AE, Huber T, Kruger P, van Gunsteren WF. The GROMOS biomolecular simulation program package. J Phys Chem A 1999;103(19):3596-3607.
39. Henriques J, Cragnell C, Skepö M. Molecular Dynamics Simulations of Intrinsically Disordered Proteins: Force Field Evaluation and Comparison with Experiment. J Chem Theory Comput 2015;11(7):3420-3431.
40. Lin Z, Van Gunsteren WF, Liu H. Conformational state-specific free energy differences by one-step perturbation: Protein secondary structure preferences of the GROMOS 43A1 and 53A6 force fields. J Comput Chem 2011;32(10):2290-2297.
41. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: A linear constraint solver for molecular simulations. J Comput Chem 1997;18(12):1463-1472.
42. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A Smooth Particle Mesh Ewald Method. J Chem Phys 1995;103(19):8577-8593.
43. Parrinello M, Rahman A. Polymorphic Transitions in Single-Crystals - a New Molecular-Dynamics Method. J Appl Phys 1981;52(12):7182-7190.
44. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. J Chem Phys 2007;126:014101.

45. Pons J, Evrard-Todeschi N, Bertho G, Gharbi-Benarous J, Sonois V, Benarous R, Girault JP. Structural studies on 24P-IkappaBalpha peptide derived from a human IkappaB-alpha protein related to the inhibition of the activity of the transcription factor NF-kappaB. Biochemistry 2007;46(11):2958-2972.
46. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: Application to microtubules and the ribosome. Proceedings of the National Academy of Sciences 2001;98(18):10037-10041.
47. Narayanan A, Jacobson MP. Computational studies of protein regulation by post-translational phosphorylation. Curr Opin Struct Biol 2009;19(2):156-163.
48. Michel Espinoza-Fonseca L, Kast D, Thomas DD. Molecular Dynamics Simulations Reveal a Disorder-to-Order Transition on Phosphorylation of Smooth Muscle Myosin. Biophysical Journal 2007;93(6):2083-2090.
49. Smart JL, McCammon JA. Phosphorylation stabilizes the N-termini of alpha-helices. Biopolymers 1999;49(3):225-233.
50. Miranda FF, Teigen K, Thorolfsson M, Svebak RM, Knappskog PM, Flatmark T, Martinez A. Phosphorylation and mutations of Ser(16) in human phenylalanine hydroxylase. Kinetic and structural effects. J Biol Chem 2002;277(43):40937-40943.
51. Dal Peraro M, Alber F, Carloni P. Ser133 phosphate-KIX interactions in the CREB-CBP complex: an ab initio molecular dynamics study. Eur Biophys J 2001;30(1):75-81.
52. Steinmetz MO, Jahnke W, Towbin H, Garcia-Echeverria C, Voshol H, Muller D, van Oostrum J. Phosphorylation disrupts the central helix in Op18/stathmin and suppresses binding to tubulin. EMBO Rep 2001;2(6):505-510.
53. Yang J, Cron P, Thompson V, Good VM, Hess D, Hemmings BA, Barford D. Molecular mechanism for the regulation of protein kinase B/Akt by hydrophobic motif phosphorylation. Mol Cell 2002;9(6):1227-1240.
54. Evrard-Todeschi N, Pons J, Gharbi-Benarous J, Bertho G, Benarous R, Girault JP. Structure of the complex between phosphorylated substrates and the SCF beta-TrCP ubiquitin ligase receptor: a combined NMR, molecular modeling, and docking approach. J Chem Inf Model 2008;48(12):2350-2361.
55. Elcock AH. Molecular simulations of diffusion and association in multimacromolecular systems. Numerical Computer Methods, Pt D 2004;383:166-198.
56. Fogolari F, Brigo A, Molinari H. The Poisson-Boltzmann equation for biomolecular electrostatics: a tool for structural biology. Journal of Molecular Recognition 2002;15(6):377-392.
57. Sheinerman FB, Norel R, Honig B. Electrostatic aspects of protein-protein interactions. Current Opinion in Structural Biology 2000;10(2):153-159.
58. Roff M, Thompson J, Rodriguez MS, Jacque J-M, Baleux F, Arenzana-Seisdedos F, Hay RT. Role of IB Ubiquitination in Signal-induced Activation of NF-B in Vivo. Journal of Biological Chemistry 1996;271(13):7844-7850.

Figure Legends

Scheme 1 Phosphorylation of IκBα by IKK leads to a degron, recognized by SCF(β-TrCP) and subsequent ubiquitination by E3. This activates the NF-κB signaling pathway whereas IκBα is degraded in the 26S proteasome.

Figure 1 (A) A ribbon representation of the structure of a complete IκBα and its phosphorylation sites bound to the partnering NF-κB, composed of the RelA and p50 subunits. RelA is colored in pink and p50 colored in brown. The ankyrin repeat domain and the C-terminal of IκBα are displayed in grey with the modeled signal receiving domain (SRD) as part of the N-terminal region colored according to its secondary structure elements with the phosphorylation sites pSer32 and pSer36 displayed in atomic detail. (B). A multiple sequence alignment of the N-terminal regulatory region, the degron motif constituted of residues DSGϕXS, with ϕ corresponding to a hydrophobic residue (here isoleucine) and X to any amino acid. (C) A close-up of the IκBα SRD displaying the modeled helical structural elements displayed in purple. The phosphorylation sites, Ser32 and Ser36, flank the degron motif highlighted in green.

Figure 2 (A) A ribbon representation of IκBα, magnifying the stretch of residues IκBα$^{28\text{-}40}$ situated on the SRD encompassing the two phosphorylation sites, Ser32 and Ser36. Root mean square fluctuation (RMSF) of the backbone atoms of segment IκBα$^{28\text{-}40}$ calculated from individual averages of three independent trajectories for each system (B) and displayed per residue (C). The highlighted regions display the sites of phosphorylation, residues 32 and 36. Root mean square deviation (RMSD) of backbone atoms for segment IκBα$^{28\text{-}40}$ during 500 ns calculated as individual averages from the three independent trajectories for each system (D) and displayed over time (E). Both RMSD and RMSF were calculated relative to the initial conformation after a least-squares fitting of IκBα$^{28\text{-}40}$. The error bars represent the

standard errors. A more stable and lower RMSD is observed in the double-phosphorylated system.

Figure 3 The $Ser32^{C\alpha}$-$Ser36^{C\alpha}$ distance compared between the different simulation systems plotted against time (A) and displayed as averages over each individual trajectory (B) with the error bars indicating the standard errors. Representative structure of the degron containing segment $I\kappa B\alpha^{31-37}$ in the unphosphorylated (C) and the double-phosphorylated system (D).

Figure 4 Relative solvent accessible surface areas (SASA) of Ser32 and Ser36 in the unphosphorylated (A), the pSer36 mono-phosphorylated (B), pSer32 mono-phosphorylated (C), and double-phosphorylated (D) systems. Ser32 and Ser36 are depicted in cyan and purple, respectively.

Figure 5 (A) Cartoon representation of double-phosphorylated $I\kappa B\alpha^{1-70}$ color coordinated according to secondary structure elements. The green shaded segment flanked with the red colored spherical phosphoserines denotes the degron in IκBα. The N-terminal, depicted in yellow, is shown to interact closely with pSer36 that forms hydrogen bonds with residues Met1 and Gln3 situated right by the N-terminal tail. Additional unique hydrogen bonds formed upon phosphorylation in the vicinity of the phosphorylation site involve residues Asp35 and Gln44 (B) and Met37 and Glu43 (C). The hydrogen bond interactions deemed significant are present in at least 30% of the simulation period in at least 2 out of 3 of the replicate runs. The structure is a snapshot of the last frame of the 500 ns simulation trajectory from the simulation run that has the lowest RMSD relative to the average structure.

Figure 6 Electrostatic potentials of the SRD mapped onto the van-der-Waals protein surface of IκBα calculated by APBS[46] in the nIκBα (A), pSer36IκBα (B), pSer32IκBα (C) and the ppIκBα states (D). The position of the site of phosphorylation is indicated by the cartoon figure of segment $I\kappa B\alpha^{31-37}$. (E) The electrostatic surface potential on the WD40 domain of β-TrCP, showing a top view of the binding interface to a double-phosphorylated IκBα.

Negative potentials of -10 kT/e are depicted in red, and positive potentials of +10 kT/e are depicted in blue. The residues making intermolecular contacts with the phosphoserines in the degron motif are indicated in their respective positions.

Scheme 2 A hypothetical model of the sequential phosphorylation of IκBα by the protein kinase IKK. Based on the simulations, Ser36 is phosphorylated first by the IKK, which in turn leads to an increased solvent accessible surface area of the second phosphorylation site, Ser32. With an enhanced exposure to the solvent, Ser32 is in turn phosphorylated by the kinase, with the effect of an extended structural loop conformation of the degron segment.