Reduction in DNA-binding affinity of Cys$_2$His$_2$ zinc finger proteins by linker phosphorylation

Derek Jantz* and Jeremy M. Berg†

Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD 21205

Communicated by Richard H. Holm, Harvard University, Cambridge, MA, March 29, 2004 (received for review December 18, 2003)

Cys$_2$His$_2$ zinc finger proteins make up the largest class of transcription factors encoded in the genomes of higher eukaryotes. Recent studies of the Ikaros transcription factor demonstrated that this zinc finger protein undergoes cell cycle-dependent changes in association with DNA that seem to be due to phosphorylation of Thr or Ser residues in the linker regions connecting adjacent zinc finger domains. The high degree of conservation of this linker sequence within the Cys$_2$His$_2$ superfamily suggested a common mechanism for the cell cycle-dependent modulation of DNA-binding affinity throughout this large class of transcription factors. The effects of linker phosphorylation on DNA-binding affinity were investigated through a direct comparison of the DNA-binding properties of four synthetic zinc finger proteins produced by native chemical ligation. The four proteins, comprising three zinc finger domains joined by two consensus Thr-Gly-Glu-Lys-Pro linkers, correspond to all four possible combinations of linker Thr phosphorylation states. Fluorescence-based DNA-binding studies of a specific DNA-binding site revealed that phosphorylation of a single linker reduced binding affinity ~40-fold, whereas phosphorylation of both linkers reduced binding affinity 130-fold. These results with purified components demonstrate that linker phosphorylation does, indeed, produce a significant reduction in DNA-binding affinity and support a model wherein a single cell cycle-dependent Ser/Thr kinase could simultaneously inactivate a large number of zinc finger transcription factors.

The Cys$_2$His$_2$ zinc finger proteins represent one of the largest classes of proteins found in the genomes of all eukaryotes that have been sequenced to date (1, 2). These proteins typically have multidomain architectures in which individual zinc-binding domains approximating the consensus sequence (Tyr,Phe)-X-Cys$_2$His$_2$. Indeed, the conservation of this sequence is typically insufficient to explain the high degree of evolutionary conservation while at the same time offering a plausible explanation for the linker’s high degree of evolutionary conservation while at the same time offering a plausible explanation for the linker’s high degree of conservation. This linker sequence acts as the substrate for an unidentified cellular kinase capable of phosphorylating Cys$_2$His$_2$ zinc finger proteins in a cell cycle-dependent manner. It has long been known that transcription ceases upon entry into mitosis and that this phenomenon is due, in part, to the mitotic inactivation of gene-specific transcription factors (14). The proposal that a cell cycle-dependent kinase can phosphorylate the Thr-Gly-Glu-Lys-Pro linker sequence provides a plausible explanation for the linker’s high degree of evolutionary conservation while at the same time offering a mechanism for the mitotic inactivation of a large number of sequence-specific transcription factors.

Because the previous report was limited to semipurified components from whole cell extracts, it was impossible for the authors to evaluate the effects of linker phosphorylation on DNA-binding affinity in a quantitative manner. To test the hypothesis that phosphorylation of linker sequences is responsible for a marked decrease in DNA-binding affinity and to provide quantitative information regarding the impact of the addition of each phosphoryl group, we compared the DNA-binding properties of a prototypical three-zinc finger protein in its four possible states of linker phosphorylation. The production of these proteins was made possible through the use of native chemical ligation (15) to join synthetic peptide fragments incorporating phosphothreonine in the linker sequence. DNA-binding studies revealed that phosphorylation of the Thr residue in a single linker produced a reduction in DNA-binding affinity of ~40-fold (at pH 7.5, 125 mM NaCl), whereas phosphorylation of both linkers produced a 130-fold reduction. These results support the hypothesis that a cellular kinase capable of phosphorylating Thr (or Ser) residues in the conserved linker in Cys$_2$His$_2$ zinc finger proteins would be able to inactivate a large number of sequence-specific transcription factors in a coordinated fashion.

Materials and Methods

Construction of Zinc Finger Proteins. The sequence of the zinc finger protein QNK-QNK-RHR is given in Fig. 1. Proteins corresponding to all four phosphorylation states of QNK-QNK-RHR were synthesized by means of the strategy outlined in Fig. 2. All peptides were made by fluorenylmethoxycarbonyl (Fmoc)-based solid phase peptide synthesis on a Milligen 9050 peptide synthesizer. A peptide comprising residues 38–65 (corresponding to the middle domain) was first synthesized on a sulfamylbutyryl AM resin (Novabiochem) to yield a C-terminal thioester as described (16). To prevent an intramolecular ligation reaction, the peptide was liberated from the resin without prior removal of the Fmoc group from the N-terminal cysteine residue. Side chain deprotection in Reagent R (95% trifluoroacetic

Abbreviation: Fmoc, fluorenylmethoxycarbonyl.

*Present address: Duke University Medical Center, Durham, NC 27705.
†To whom correspondence should be addressed. E-mail: bergj@mail.nih.gov.

© 2004 by The National Academy of Sciences of the USA
was cleaved from the resin and deprotected in Reagent R. The domains were synthesized on a PEG-PS resin (Applied Biosystems), comprising residues 66–86 of the Thr residues that are the sites of phosphorylation and shaded. The position of the linkers formed by peptide ligation is indicated by dashed lines.

DNA-Binding Assay. All oligonucleotides were ordered from Operon Technologies (Alameda, CA). A fluorescein-labeled DNA probe was prepared by annealing oligonucleotides with the sequences 5’-CGATCGTTCAGGGCAGAACCCTGATCA-3’ and 5’-TGATCAGGXTCTGCTCAAGCATCG-3’, where X is fluorescein-5-T (protein-binding site is in italics). Three milliliters of a 5 mM solution of this probe in 20 mM Tris, pH 7.5/125 mM NaCl/10 μg/ml BSA were added to a 4.5 ml of 4.5 mM fluorescein Cu(II) fluorophore in the T format with the excitation monochromator set at 490 nm and the emission monochromator at 515 nm. Subsequent anisotropy measurements were taken after the stepwise addition of purified zinc finger protein, allowing 5 min at room temperature before each measurement. All titrations were carried out in triplicate. No significant changes in fluorescence intensity were observed over the course of the titrations. Plots of the change in anisotropy as a function of added protein concentration were fit by using KALEIDAGRAPH to yield dissociation constants for the protein–DNA complexes. To rule out the possibility that binding was being influenced by the fluorophore on the labeled DNA probe, competition experiments were conducted to determine dissociation constants for the four proteins in complex with an unlabeled oligonucleotide binding site (data not shown). In all cases, Kd values determined for the labeled and unlabeled sites were within experimental error.

Alkaline Phosphatase Treatment. A 100-μmol solution of each protein was prepared in 20 mM Tris, pH 8.0, with 100 mM NaCl, 10 mM MgCl2, 1 mM DTT, and 500 μM ZnCl2. Two microliters of calf intestinal phosphatase (20 units, New England Biolabs) were added to 50 μl of each protein solution, and the reactions were heated at 37°C overnight. These solutions were then diluted 10-fold into 20 mM Tris, pH 7.5, with 125 mM NaCl and used fluorescence anisotropy-based DNA-binding assays.

Results

Synthesis of Phosphoproteins. The four proteins under investigation are variants of a designed zinc finger protein termed QNK-QNK-RHR (referred to as QNK) showing the numbering scheme. Metal-binding domains are shown as cylinders, and the linkers are shown as thick lines.

Fig. 1.  The amino acid sequence of the three-zinc finger protein QNK-QNK-RHR (referred to as QNK) showing the numbering scheme. The Thr residues that are the sites of phosphorylation and shaded. The positions of the linkers formed by peptide ligation are indicated by dashed lines.

Fig. 2.  The synthetic scheme for the preparation of QNK-QNK-RHR in its fully phosphorylated form (referred to as QNKp123). The zinc finger domains are shown as cylinders, and the linkers are shown as thick lines.
Characterization of the synthetic zinc finger proteins. (A) SDS/PAGE gel showing the four purified proteins. Molecular masses (m + 1) of 9,780.4 daltons (9,778.5 calculated) (QNk), 9,860.6 daltons (9,858.5 calculated) (QNk12), 9,860.9 daltons (9,858.5 calculated), and 9,941.3 daltons (9,938.5 calculated) (QNk123) were determined by matrix-assisted laser desorption ionization mass spectrometry. (B) Cobalt titrations to determine metal binding affinity. All four proteins bound three equivalents of cobalt each, and titration data were fit to yield dissociation constants lower than 1 × 10⁻⁷ M for all metal binding sites.

Fig. 3. Characterization of the synthetic zinc finger proteins. (A) SDS/PAGE gel showing the four purified proteins. Molecular masses (m + 1) of 9,780.4 daltons (9,778.5 calculated) (QNk), 9,860.6 daltons (9,858.5 calculated) (QNk12), 9,860.9 daltons (9,858.5 calculated), and 9,941.3 daltons (9,938.5 calculated) (QNk123) were determined by matrix-assisted laser desorption ionization mass spectrometry. (B) Cobalt titrations to determine metal binding affinity. All four proteins bound three equivalents of cobalt each, and titration data were fit to yield dissociation constants lower than 1 × 10⁻⁷ M for all metal binding sites.

DNA-Binding Activities of the Phosphoproteins. A fluorescence anisotropy-based DNA binding assay was used to evaluate the affinities of the four proteins for DNA (Fig. 4). The QNK-QNK-RHR protein, named for the DNA contacting amino acids in positions 1, 3, and 6 of its three zinc finger helices, binds with high affinity to the DNA site 5′-GA/GGAAAGAA-3′ (22). The protein is known to be highly tolerant of substitutions in the fifth and eighth positions of its binding site so that a fluorescein-labeled, double-stranded DNA probe containing the sequence 5′-GAGGCAGAA-3′ was suitable for the characterization of DNA binding. The results of these binding studies are shown in Fig. 3. The unphosphorylated QNK protein was found to bind this sequence with a Kd value of 1.7 ± 0.2 nM. In contrast, QNKp12 bound the DNA probe with a Kd value of 62 ± 6 nM, a 36-fold reduction in affinity. QNKp23 bound the DNA site with a Kd value of 78 ± 9 nM, a 46-fold loss in affinity over the unphosphorylated protein. The doubly phosphorylated QNKp123 bound to the DNA with still lower affinity, binding to the fluorescent probe with a Kd value of 221 ± 8 nM, a 130-fold reduction in affinity relative to QNK.

DNA-binding activity was readily restored to the phosphorylated proteins by treatment with alkaline phosphatase (Fig. 3).

Discussion

It was observed over 40 years ago that the passage of eukaryotic cells into mitosis is accompanied by a pause in transcription until cell division is complete (23). There seem to be several underlying causes for this phenomenon, including limited access to the DNA template upon chromatin condensation, the production of
general transcription repressor proteins, and the inactivation of transcription machinery. The latter mechanism has both a general and a gene-specific component. A general block to transcription is afforded by the inhibition of RNA polymerase II through phosphorylation of transcription factors TFIIID (24) and TFIIH (25) and of the polymerase itself (26). It has also been observed that several components of the SWI/SNF chromatin remodeling complex are phosphorylated upon entry into mitosis and excluded from mitotic chromosomes (27). In addition to these general mechanisms for inhibiting transcription, several gene-specific transcription factors are known to be inactivated upon entry into mitosis. The oncoproteins Myc and Myb were found to be hypophosphorylated at mitosis (27) as were several members of the POU family of homeodomain proteins (27–30). These site-specific DNA-binding proteins all exhibited reduced DNA-binding activity in mitotic cell extracts. It is unclear why these gene-specific transcription factors are inactivated upon entry into mitosis, but it has been proposed that this type of inhibition is necessary for cell division and the proper establishment of expression patterns upon completion of mitosis (13).

The existence of a common mechanism for the mitotic inactivation of Cys2His2 zinc finger transcription factors is an attractive hypothesis given the abundance of this type of protein encoded in eukaryotic proteomes. The simplest model for such a system would be the direct inhibition of DNA binding through phosphorylation of a conserved site within all zinc finger domains by a cell cycle-dependent kinase. This model places strict restrictions on the site of phosphorylation within the zinc finger region. The site must be (i) conserved between proteins and, therefore, not directly involved in site-specific DNA recognition, (ii) accessible to the kinase active site when bound to DNA, and (iii) able to interfere with DNA binding when phosphorylated. The Thr-Gly-Glu-Lys-Pro linker sequence appears to meet these criteria. It is highly conserved between zinc finger proteins but is not part of the DNA-recognition helix responsible for conferring sequence specificity. As such, the linker sequence can remain constant even as members of the family evolve to have novel DNA recognition properties. The accessibility of the site to a cellular kinase is difficult to investigate until the kinase in question is identified. The existing structures of zinc finger proteins in complex with DNA, however, reveal that the linker forms an extended structure that is reasonably distant from the DNA. The threonine side chain that is proposed to be the target for phosphorylation is on the outer surface of this loop and exposed to solvent. As such, this residue is positioned such that it would be accessible to regulatory factors even when bound to DNA. Also consistent with this proposal is our finding that calf-intestinal phosphatase was able to remove the phosphate groups from our phosphorylated proteins. This result indicates that, at least in the absence of DNA, the zinc finger structure does not interfere substantially with enzymatic activity at the linker Thr residue. Moreover, this result lends support to the proposal that a cellular phosphatase could reactivate DNA binding and reestablish transcription patterns following mitosis. Our results demonstrate conclusively that phosphorylation of the linker Thr residues does, indeed, lead to a substantial decrease in DNA-binding affinity for a prototypical zinc finger-based DNA-binding protein. The biophysical explanation for the decrease in DNA-binding affinity remains to be established. Examination of the structure of the unphosphorylated protein bound to DNA (18) reveals several features: (i) there is sufficient space for a phosphoryl group to be introduced without the introduction of unfavorable steric clashes with either the protein backbone or the DNA; (ii) if no structure changes take place, the phosphoryl group would lie ~10 Å from two DNA backbone phosphate groups; and (iii) the introduced phosphoryl group could interact with side chains within the linker or the zinc finger preceding it. Based on these features, several mechanisms may contribute to the decrease in DNA binding. Direct electrostatic interactions between the introduced phosphoryl group and the negatively charged DNA backbone could be responsible; an interaction between singly negatively charged groups separated by 10 Å in water corresponds to 0.4 kcal/mol (1 cal = 4.184 J), a value expected to be reduced further by the presence of counterions. The interaction of a single phosphate on the protein with two DNA backbone phosphates would be expected to decrease the affinity by approximately a factor of 4. This value is well below the 40-fold reduction in binding affinity that we observed with the singly phosphorylated proteins, suggesting that additional factors are at work. It is also possible that linker phosphorylation inhibits DNA binding by disrupting structural elements within the protein itself. It has been observed that the linker region becomes structured upon DNA binding, and this structure has been proposed to contribute to the relative positioning of individual zinc finger domains (4–6). In particular, the conserved threonine seems to act as part of a C-cap, which terminates the DNA-recognition helix of the preceding zinc finger domain (7). Phosphorylation of the Thr side chain might disrupt these contacts and destabilize linker structures appropriate for positioning adjacent zinc finger domains along the DNA. Alternatively, the phosphorylation might stabilize difference conformations of the linker that are well suited for DNA binding. A potential explanation that phosphorylation leads to zinc release and, hence, destabilization of the zinc finger structure is ruled out by our observation that all four proteins exhibit comparable affinities for metal ions as demonstrated by direct cobalt(II)-binding studies.

In summary, we have used solid-phase peptide synthesis techniques to produce a set of zinc finger proteins that have precisely incorporated phosphoryl groups and have used these proteins to correlate phosphorylation state with DNA-binding affinity. Through the direct demonstration that linker phosphorylation does, in fact, impair DNA binding, we provide support for the existence of a cellular kinase/phosphatase pair capable of inactivating/reactivating Cys2His2 zinc finger proteins in a cell cycle-dependent manner. This hypothesis provides considerable insight into the previously enigmatic observation that the linker between zinc finger domains is one of the most conserved features among members of this large and important family of proteins.

We thank Dr. Gregory J. Gatto, Jr., for useful discussions and comments on the manuscript. This work was supported by National Institute of General Medical Sciences Grant GM046257.