Gene Regulation: A Tandem SH2 Domain in Transcription Elongation Factor Spt6 Binds the Phosphorylated RNA Polymerase II C-terminal Repeat Domain (CTD)

Mai Sun, Laurent Larivièere, Stefan Dengl, Andreas Mayer and Patrick Cramer

doi: 10.1074/jbc.M110.144568 originally published online October 6, 2010

Access the most updated version of this article at doi: 10.1074/jbc.M110.144568

Find articles, minireviews, Reflections and Classics on similar topics on the JBC Affinity Sites.

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 43 references, 20 of which can be accessed free at http://www.jbc.org/content/285/53/41597.full.html#ref-list-1
Spt6 is an essential transcription elongation factor and histone chaperone that binds the C-terminal repeat domain (CTD) of RNA polymerase II. We show here that Spt6 contains a tandem SH2 domain with a novel structure and CTD-binding mode. The tandem SH2 domain binds to a serine 2-phosphorylated CTD peptide in vitro, whereas its N-terminal SH2 subdomain, which we previously characterized, does not. CTD binding requires a positively charged crevice in the C-terminal SH2 subdomain, which lacks the canonical phospho-binding pocket of SH2 domains and had previously escaped detection. The tandem SH2 domain is apparently required for transcription elongation in vivo as its deletion in cells is lethal in the presence of 6-azauracil.

Spt6 is an essential nuclear protein (1, 2) that is required for transcription elongation together with Spt4 and Spt5 (3, 4). Spt6 co-localizes with Spt5 on transcribed genes in Drosophila polytene chromosomes (5, 6). Spt6 stimulates the elongation rate of RNA polymerase II in vivo (7). There is strong evidence that Spt6 is a histone chaperone. In cells lacking functional Spt6, transcription can start from cryptic promoters within the coding regions of genes (8), and transcription reinitiation from regular promoters is facilitated (9). This suggests that Spt6 is required for the correct positioning of nucleosomes in transcribed regions. Consistently, Spt6 interacts with histone H3 and can deposit nucleosomes on DNA (10). Spt6 is further involved in mRNA 3′-end formation (11) and export (12).

The C-terminal region of mouse Spt6 binds the Ser-2-phosphorylated CTD, suggesting a mechanism for how Spt6 is recruited to elongating RNA polymerase II (12). The Spt6 C-terminal region was predicted to contain a novel CTD-binding protein module and indicate that the structure and CTD binding function of the Spt6 C-terminal region is conserved in eukaryotes.

**EXPERIMENTAL PROCEDURES**

Preparation of the Spt6 Tandem SH2 Domain—DNA containing the tandem SH2 domain of C. glabrata Spt6 (residues 1250–1444) was amplified by polymerase chain reaction from genomic DNA and was cloned into expression vector pET28b(+) (Novagen) using restriction sites NdeI and NotI. This results in an open reading frame that encodes the tandem SH2 domain with an N-terminal hexahistidine tag with a thrombin cleavage site. The sequence was modified by the introduction of codons for methionine residues at the positions of residues Leu-1309, Leu-1317, and Leu-1411 by the rapid PCR site-directed mutagenesis (17). The resulting variant was used for selenomethionine labeling as described (18, 19). The tandem SH2 domain was expressed overnight at 37 °C with a thrombin cleavage site. The sequence was modified by the introduction of codons for methionine residues at the positions of residues Leu-1309, Leu-1317, and Leu-1411 by the rapid PCR site-directed mutagenesis (17). The resulting variant was used for selenomethionine labeling as described (18, 19). The tandem SH2 domain was expressed overnight at 18 °C in Escherichia coli BL21 CodonPlus(DE3)RIIL strain cells (Stratagene). For selenomethionine labeling, we used E. coli strain B834(DE3)pLysS (Novagen). Cells were harvested, resuspended in 50 mM HEPES, pH 7.0, 1 mM NaCl, 10 mM β-mer-

---

1. To whom correspondence should be addressed. Tel.: 49-89-2180-76951; Fax: 49-89-2180-76999; E-mail: cramer@LMB.uni-muenchen.de.
2. The abbreviations used are: CTD, C-terminal repeat domain; 6-AU, 6-azauracil.
Tandem SH2 Domain in Spt6 Binds RNA Polymerase II CTD

captoethanol, 10 mM imidazole, and lysed by sonication. Cell lysates were subjected to affinity chromatography on a nickel-nitrilotriacetic acid column (Qiagen) followed by cleavage of the hexahistidine tag with thrombin and dialysis overnight at 4 °C. The tag and the undisgusted protein were removed on a second nickel-nitrilotriacetic acid column. DNA was removed by cation exchange chromatography (Mono S 10/100 GL, Amersham Biosciences). After gel filtration (Superose 12 10/300 GL, Amersham Biosciences), the sample was concentrated to 14.5 mg/ml and stored at 4 °C in 50 mM HEPES, pH 7.0, 100 mM NaCl, 5 mM DTT.

Crystal Structure Determination—The purified native and selenomethionine-labeled Spt6 tandem SH2 domain was crystallized by vapor diffusion using 22–24% PEG 8000, 50 mM MES, pH 6.5–7, 0.2 M sodium acetate as reservoir solution. Complete diffraction data at 1.7 and 1.6 Å resolution were obtained from crystals of the selenomethionine-labeled protein and unlabeled protein, respectively, and were processed by XDS and scaled by XSCE (20). The structure was solved by multilwavelength anomalous diffraction using SOLVE, and density modification was performed with RESOLVE (21). The model was automatically built by the program ARP/wARP (22) with some manual rebuilding using Coot (23) and refined with program PHENIX (24). We added water molecules initially with Coot when a peak height greater than 0.5 \( \text{R}_{\text{free}} \) was obtained. We added more water molecules manually and further refined the model with PHENIX using bulk solvent scaling and individual isotropic B-factors. Refinement statistics are shown in Table 1.

Phosphopeptide Interaction Assay—For fluorescence anisotropy measurements, we used a HORIBA Jobin-Yvon FluoroMax 3 spectrophuorometer. The CTD phospho-peptides were synthesized by PANATecs GmbH (Tübingen), had the sequence SYSPTSPSYSPTSPS (pS, phospho-serine), and were N-terminally labeled with fluorescein aminocaproic acid. The phospho-peptide (500 nm) was dissolved in 20 mM HEPES, pH 8.0, 10 mM NaCl, 5 mM DTT and was titrated at 20 °C with pure tandem SH2 domain or its mutant variants R1281K, K1434R, or R1281K/K1434R, and the fluorescence anisotropy was measured after 2 min when equilibrium was reached. Data were analyzed by non-linear least squared regression against \( A = A_0 + \Delta A \times 0.5 \times (b - (b^2 - P \times L)^{1/2}) \) with \( b = P + L + K_2 \) (\( K_2 \), apparent dissociation constant; \( P \), protein concentration; \( L \), ligand concentration).

Yeast Strains—Saccharomyces cerevisiae strains containing a C-terminal tandem affinity purification-tagged version of Spt6 and the same strain lacking amino acid residues 1250–1451 of Spt6 were as described (13).

RESULTS

Spt6 Contains a Tandem SH2 Domain—When we examined the Spt6 amino acid sequences flanking the known SH2 domain, we noted additional predicted secondary structure elements C-terminal to the known SH2 domain. In particular, the arrangement of predicted \( \alpha \)-helices and \( \beta \)-strands suggested that a second SH2 domain could be present that had escaped detection by standard domain prediction procedures. To investigate this, we prepared a recombinant variant of the Spt6 C-terminal region that spanned the known SH2 domain and the predicted SH2 domain (see “Experimental Procedures”). Soluble protein was purified and crystallized. The x-ray structure was determined by multilwavelength anomalous diffraction and refined at 1.6 Å resolution (Table 1).

The structure revealed the previously reported SH2 domain (13) and additionally a second SH2 domain (Fig. 1). The two domains, designated hereafter subdomains SH2-N and SH2-C, intimately pack against each other. Therefore, we refer to this structure as tandem SH2 domain. The C-terminal helix of subdomain SH2-N is composed of two parts (\( \alpha B1 \) and \( \alpha B2 \)). The \( \alpha B2 \) part extends into subdomain SH2-C and makes extensive hydrophobic contacts with it, thereby stabilizing the tandem structure. Superposition of the two subdomains revealed that both have a similar structure, although the N- and C-terminal helices in SH2-C were longer and shorter, respectively, than in SH2-N (Fig. 1C). Because hydrophobic core residues are conserved from yeast to human (Fig. 1A, triangles), the tandem SH2 domain fold is conserved among eukaryotes.

Unique Structure of the Tandem SH2 Domain—An arrangement of two subsequent SH2 domains was previously observed in signaling proteins present in higher eukaryotes. Structures of four proteins with two adjacent SH2 domains were published from the related tyrosine kinases ZAP-70 (25, 26) and Syk (27) and from the related tyrosine phosphatases SHP-2 (28, 29) and SHP-1 (30). A comparison revealed that these structures are all strikingly different from the Spt6 tandem SH2 domain and showed different arrangements of the two SH2 domains (Fig. 2). In addition,

![Table 1 X-ray diffraction data and refinement statistics for the Spt6 tandem SH2 domain](http://www.jbc.org/content/journal/jbc/285/53/41598.full.pdf)
the two SH2 domains are always linked by an insertion of varying structure that does not exist in the Spt6 tandem SH2 domain. Although the SH2 domains pack against each other in ZAP-70 and Syk, they do not in the SHP phosphatases and are instead positioned via additional domains present in these proteins.

**The Tandem SH2 Domain Binds to the Ser-2-phosphorylated CTD in Vitro**—We used fluorescence anisotropy to investigate whether the tandem SH2 domain binds to CTD peptides with two heptapeptide repeats (see “Experimental Procedures”). We used four peptides. One peptide was phosphorylated on both Ser-2 residues (P1), whereas another...
Tandem SH2 Domain in Spt6 Binds RNA Polymerase II CTD

FIGURE 2. Comparison with other SH2 domain tandem arrangements. The SH2 domains of ZAP-70 (Protein Data Bank (PDB) code 2oq1), SHP-1 (PDB code 2b3o), and SHP-2 (PDB code 2shp) are shown, with their N-terminal and C-terminal SH2 domains in blue and magenta, respectively. Additional structural elements between both SH2 domains are represented in orange, and a bound phospho-peptide in ZAP-70 is in red. The N-terminal SH2 domains in these proteins were superimposed onto Spt6 subdomain SH2-N.

One was unphosphorylated (P4). One peptide was phosphorylated only on the N-terminal Ser-2 residue (P2), and one was phosphorylated only at the C-terminal Ser-2 residue (P3). Titration equilibrium measurements revealed binding curves indicative for a specific protein-peptide interaction for the Ser-2-phosphorylated peptides P1, P2, and P3, whereas the unphosphorylated peptide P4 did not bind (Fig. 3A and Table 2). Regression analysis led to a dissociation constant of 10 μM for P1, and the binding affinity dropped 3–5-fold when only a single phosphorylated Ser-2 was present. Isolated SH2-N did not show specific binding to the same phospho-peptide (13). Thus, the tandem SH2 domain, but not SH2-N alone, binds to Ser-2-phosphorylated CTD peptides in vitro, and phosphorylation of two residues enhances binding.

CTD Binding Requires a Conserved Non-canonical Surface Crevice—The structure revealed that the canonical phospho-binding pocket in SH2-N remains accessible in the tandem domain (Fig. 3B, Pocket 1). The region in SH2-C corresponding to the canonical pocket 1, however, lacks the conserved arginine found in all other SH2 domains, and instead the side chain of Tyr-1393 occupies the space of the canonical pocket. However, a conserved surface crevice was observed on the opposite side of SH2-C (Fig. 3B). This crevice is positively charged and is flanked by the invariant residue Lys-1434 that could contribute to phospho-peptide binding (Figs. 1A and 3, B and C).

To investigate whether SH2-N pocket 1 and/or the SH2-C crevice are required for phospho-CTD binding, we prepared three mutant variants of the tandem SH2 domain and tested them for binding the doubly Ser-2-phosphorylated peptide P1 (Fig. 3A). Mutation of residue Arg-1281 in pocket 1 of SH2-N to lysine led to a 1.4-fold increase of the dissociation constant. Mutation of residue Lys-1434 in the crevice of SH2-C to alanine caused a 2.5-fold increase of the dissociation constant. When both point mutations were combined into a double mutant variant of the tandem SH2 domain, binding affinity dropped strongly. Binding to the phospho-CTD peptide did not reach saturation and could not be fitted using a simple binding model. This non-saturation behavior shows that the double mutant only interacts non-specifically with the phospho-peptide. Thus, both patches contribute to the binding affinity, and the CTD peptide apparently extends over the tandem domain surface from pocket 1 in SH2-N via a positively charged surface to the non-canonical crevice in SH2-C that strongly contributes to the binding affinity (Fig. 3).

We next investigated how the domain mutations influence binding of CTD peptides with single Ser-2 phosphorylations. The mutant tandem SH2 domain variants generally did not bind to single phosphorylated CTD peptides anymore, except for variant R1281K, which bound the single phosphorylated CTD peptide P2 with almost the same affinity as the wild-type domain. This indicated that the N-terminal phosphorylated Ser-2 residue binds to the conserved surface crevice on SH2-C. However, confirmation of this proposal and additional details of the interaction await a structure of the tandem domain in complex with a CTD peptide. Modeling indicates that a single CTD repeat cannot span the distance between the binding sites on the two SH2 subdomains, consistent with multiple CTD repeats contributing in the interaction.

Deletion of the Tandem SH2 Domain Causes an Elongation Defect in Vivo—Deletion of the C-terminal region in Spt6 that corresponds to the tandem SH2 domain in yeast causes a slow growth phenotype (13, 31). To investigate whether this phenotype goes along with a transcription elongation defect, we tested the sensitivity of this deletion strain to 6-AU. Sensitivity to this nucleotide-depleting drug is observed in many elongation factor mutants (32), including mutants of the Spt6 gene (4, 33). We observed that the deletion strain failed to grow on 6-AU-containing (50 μg/ml) medium (Fig. 4). This shows that the Spt6 tandem SH2 domain is essential for cell viability under conditions when transcription elongation is impaired.
Here we report that the C-terminal region of the essential transcription elongation factor Spt6 contains a tandem SH2 domain that combines the previously described SH2 domain (subdomain SH2-N) with a second, subsequent SH2 domain (subdomain SH2-C) into a single domain of novel structure. We show that the tandem SH2 domain, in contrast to SH2-N alone, binds to Ser-2-phosphorylated CTD peptides in vitro.

**FIGURE 3.** The tandem SH2 domain binds the RNA polymerase II CTD in vitro. A, fluorescence anisotropy titration of the CTD-derived phospho-peptide P1 (SYpSPTSPSYpSPTSP; pS, phospho-serine) with different Spt6 tandem SH2 domain variants (open circles, wild-type tandem SH2 domain; open triangles, variant R1281K; filled circles, variant K1434A; filled triangles, variant R1281K/K1434A). The titration with the variant R1281K/K1434A could not be fitted with a simple binding model. The solid lines represent non-linear least square fits. B, Spt6 tandem SH2 domain surface conservation. Invariant and conserved residues are colored according to Fig. 1A. Indicated with circles are pocket 1 in SH2-N, the region that would correspond to pocket 1 in SH2-C, and the CTD-binding crevice in SH2-C. Residues Arg-1281 and Lys-1434 are represented as sticks and are labeled. The left and right views are related by a 120° rotation around the vertical axis. C, surface charge distribution. The domain surface is color-coded according to the electrostatic potential from negative (red) to positive (blue). Electrostatic surface potentials were calculated with APBS. The left and right views are related by a 120° rotation around the vertical axis.

**TABLE 2**

Dissociation constants (mM) obtained by fluorescence anisotropy

<table>
<thead>
<tr>
<th>CTD peptide</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tandem SH2 variant</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>10.4 ± 0.2</td>
<td>34.7 ± 2.2</td>
<td>46.0 ± 2.3</td>
<td>NF</td>
</tr>
<tr>
<td>R1281K</td>
<td>13.9 ± 0.6</td>
<td>40.7 ± 1.1</td>
<td>NF</td>
<td>ND</td>
</tr>
<tr>
<td>K1434A</td>
<td>25.7 ± 0.7</td>
<td>NF</td>
<td>NF</td>
<td>ND</td>
</tr>
<tr>
<td>R1281K/K1434A</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
<td>NF</td>
</tr>
</tbody>
</table>

**FIGURE 4.** The SH2 tandem domain is apparently involved in transcription elongation in vivo. Cultures of wild-type and Spt6 tandem domain deletion yeast strains (13, 31) were grown in YPD medium at 30 °C overnight and diluted to an optical density at 600 nm of 5 with fresh medium. The same amount of cells was spotted on plates in 10-fold serial dilutions. Plates were incubated for 5 days at 30 °C and inspected daily. 6-AU was added to SD-Ura plates at a concentration of 50 µg/ml.
Tandem SH2 Domain in Spt6 Binds RNA Polymerase II CTD

A single phosphorylated Ser-2 residue is required and sufficient for binding, but the presence of a second phosphorylated Ser-2 residue strongly enhances binding, consistent with published data (34–36). The tandem domain is also required for cell growth in the presence of the nucleotide-depleting drug 6-AU in vivo, indicating its importance for transcription elongation.

We further show that the tandem SH2 domain of Spt6 is unique in its structure and mode of CTD binding. Two adjacent SH2 domains are found in several proteins in higher eukaryotes but are arranged in a different way in known three-dimensional structures (Fig. 2). Also, the tandem domain contains a canonical phospho-binding pocket only in its subdomain SH2-N, but not in its subdomain SH2-C, which lacks the invariant arginine residue found in all other SH2 domains. Consistent with observed structural differences, the Spt6 tandem SH2 domain has a unique phospho-peptide binding mode. Two surface patches are involved in CTD binding, the canonical phospho-binding pocket in subdomain SH2-N and a non-canonical basic crevice in subdomain SH2-C that comprises an invariant lysine residue that contributes to binding. Mutating either one of these two conserved basic residues on the two CTD-binding surface patches resulted in a weaker interaction. It is likely that the SH2-C crevice interacts with a phospho-serine residue because a CTD peptide with a single phosphorylated Ser-2 residue binds equally well to wild-type tandem SH2 domain and a variant in which the SH2-N pocket is mutated. These results are consistent with an extended peptide-binding surface that reaches from the SH2-N pocket to the SH2-C crevice. Structural features that may provide specificity for CTD binding are, however, difficult to identify in the absence of a complex crystal structure.

Canonical SH2 domains bind phosphorylated target peptides via two pockets, one that recognizes the phosphorylated tyrosine side chain (pocket 1) and one that contacts residues C-terminal of the phospho-tyrosine (pocket 2), generally at register +3. Pocket 2 generally contributes to the specificity of the interaction. In Spt6 SH2-N, pocket 1 is present and used for peptide binding, but the previously described pocket 2, which was proposed to play a similar role as in other SH2 domains (13), is occluded in the tandem SH2 domain structure (Fig. 3B). In SH2-C, neither pocket 1 nor pocket 2 is observed, but instead a conserved crevice on the opposite face contributes to peptide binding.

Despite all these differences between the Spt6 tandem SH2 domain and other arrangements of two SH2 domains, there is a functional analogy. Consistent with our observations, individual ZAP-70 SH2 domains do not bind the target phosphopeptide but do so when in a tandem arrangement (37). More generally, the phospho-peptide binding specificity of single SH2 domains is increased by the simultaneous binding of the peptide to an additional domain (38). Tandem SH2 domains therefore may generally increase target specificity due to cooperative peptide binding as suggested (39).

Recent data indicate that the tandem SH2 domain is required for retaining Spt6 at the end of transcribed genes, rather than recruiting it to genes. We have recently shown by high resolution genome-wide chromatin immunoprecipitation that Spt6 is recruited during an early transition in the 5′-region of genes near the transcription start site before high levels of Ser-2 phosphorylation are reached (40). Deletion of the genomic region encoding the tandem SH2 domain lowered the level of recruitment, in particular at the 3′-region around the polyadenylation site, where levels of Ser-2 phosphorylation are high (40). These results indicated that Spt6 is recruited near the transcription start site with the use of a CTD Ser-2 phosphorylation-independent mechanism in vivo, but full recruitment and retaining of Spt6 at the polyadenylation site require the CTD-binding tandem SH2 domain.

Acknowledgments—We thank Dirk Kostrewa and other members of the Cramer laboratory for help. Part of this work was performed at the Swiss Light Source (SLS) at the Paul Scherrer Institute (Villigen, Switzerland). We thank the crystalization facility at the Max Planck Institute (MPI) for Biochemistry, Martinsried, Germany.

REFERENCES


